Relativistic Quantum Mechanics and Quantum Field Theory

V. Devanathan

Alpha
Science

Relativistic Quantum Mechanics and Quantum Field Theory

Relativistic Quantum Mechanics and Quantum Field Theory

V. Devanathan

Alpha Science International Ltd.
Oxford, U.K.

Relativistic Quantum Mechanics and Quantum Field Theory
272 pgs. | 27 figs. | 4 tbls.

V. Devanathan
Professor of Nuclear Physics (Retd.)
University of Madras
Guindy Campus
Chennai, India

Copyright © 2011

ALPHA SCIENCE INTERNATIONAL LTD.
7200 The Quorum, Oxford Business Park North
Garsington Road, Oxford OX4 2JZ, U.K.

www.alphasci.com

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the publisher.

Printed from the camera-ready copy provided by the Author.

ISBN 978-1-84265-657-0

Printed in India

To my esteemed colleague

Dr. M. Anandakrishnan

who has contributed greatly
to the development of higher education in India

Preface

Ever since I wrote my book on Quantum Mechanics, I wished to follow it up with another book on Relativistic Quantum Mechanics and Quantum Field Theory as a sequel, since most of the Universities have a two-semester course in M.Sc. (Physics), one on Quantum Mechanics and the other on Relativistic Quantum Mechanics and Quantum Field Theory.

In the meantime, the book on "Relativistic Quantum Mechanics" by R. Parthasarathy has come out in addition to the already published work "The First Book on Quantum Field Theory" by Lahri and Pal from the present publisher. This has made me more hesitant to go ahead with the project and hence the delay in completing the manuscript of the book. Hence one may rightly ask, what is the need for another book when there are so many excellent books on the subject?

In Chennai, December-January is the season for music festivals. There are very many music concert halls, to which the people flock together to listen to the musicians of their choice. Most of the musicians sing almost the same songs but with different flavours. The discerning public not only enjoy the different flavours and the small variations in style but welcome the variety in unison and the music connoisseurs wax eloquent on the unity in diversity.

Thus each author has his own style of presentation and a distinct approach to the subject. In the same way, I have developed my own style of presentation and a distinctive approach to the subject. A discerning student will find some topics in one book more intelligible than in other books and he would wish to refer to more books to have a holistic view of the subject.

In that spirit, I have gathered courage to venture on the publication of this book. I have been greatly fascinated by Feynman's intuitive approach to Quantum Electrodynamics which has given a visual picture of the processes that we study and a thumb rule for calculating the matrix elements. As Schwinger once remarked, Feynman has indeed carried the

Quantum Electrodynamics to the "masses". More than Feynman, it was Dyson who had popularized Feynman diagrams and demonstrated that Feynman's intuitive approach is equivalent to the more general field theoretic approach of Tomonago and Schwinger. You will find in this book the highlights of Feynman's contribution and how it leads to the same results that one can arrive by the Quantum Field Theory.

Chapters 1 to 5 deal with single particle relativistic wave equations and Feynman's approach to Quantum Electrodynamics. Chapters 6 to 10 deal with the elements of Quantum Field Theory and how one can apply it to the study of Quantum Electrodynamics. It is shown that Quantum Field Theory is more general and applicable not only to electromagnetic interaction but has the potential of being extended to weak interaction and strong interaction. Thus the Quantum Field Theory has unified the three of the four known fundamental interactions, electromagnetic, weak and strong interactions. Only the gravitational interaction lies outside its purview.

It is being felt that the Quantum Field Theory should be the low-energy limit (less than 1000 GeV) of a more fundamental theory. String Theory and String Theory with super-symmetry known as Superstring theory seem to be the potential candidates for such a more fundamental theory which is applicable to all the four fundamental interactions. The last chapter on Recent Developments deals with the emerging scenario on this subject.

The author has not made any serious effort to include all the references. Some original papers have been cited in the text as footnote and some books and review articles are included in the Bibliography at the end.

A notable feature of this book is the inclusion of review questions, problems and solutions to problems at the end of each chapter which will definitely promote a clearer and deeper understanding and appreciation of the subject. The author welcomes any suggestion for improvement and he can be contacted by email: vdevanathan@hotmail.com.

One cannot expect the reader to start from the first page and go to the last page in a sequential order. Notations and symbols are defined usually in the text when they are first introduced. The difficulty that a reader usually faces is in deciphering notations and symbols when he wants to refer to any specific topic. To help such a reader, I have appended Appendix A: List of Symbols and Notations.

For the past several years, I had the privilege of organizing the Summer

Training Programmes in Physics (STPIP) for the post-graduate students in Physics, under the joint auspices of the University of Madras, Tamil Nadu Academy of Sciences and the Science City. This has enabled me to interact with the present and past students and the faculties and sustain my interest in bringing out this book. Thanks are due to Dr. M. Anandakrishnan, Dr. P. Iyamperumal, Dr. J. Devasenapathy, Dr. V. Ravichandran, Dr. A. Stephen and all the members of the Department of Nuclear Physics for extending their fullest cooperation to me in organizing these summer programs.

I am grateful to Dr. K. Sivaji for making available all the relevant literature on the subject and to Mr. S. Gopinath for offering the software support. Special thanks are due to Mr. N. K. Mehra, Managing Director of Narosa Publishing House for readily agreeing to undertake the publication of this book and tolerating the inordinate delay in my submission of the camera-ready manuscript.

In conclusion, I wish to thank the members of my family for their tolerance and support that they have extended to me in completing this project.

V. Devanathan
March 2011

Contents

Preface **vii**

1 Introduction **1**

1.1 The Schrödinger Equation 2
1.2 The Klein-Gordon Equation 3
1.3 The Dirac Equation 5
1.4 Feynman's Positron Theory 11
1.5 Quantum Electrodynamics 12
1.6 Quantum Field Theory 13
Review Questions 16
Problems 16
Solutions to Problems 16

2 The Dirac Equation and its Solutions **21**

2.1 The Free Particle Solutions 21
2.2 Orthogonal and Closure Properties 24
2.3 Projection Operators 26
2.4 Sum Over Spin States 26
2.5 In Feynman's Notation 27
2.5.1 Positive eigenvalue states 28
2.5.2 Negative eigenvalue states 29
2.5.3 Sum over spin states 33
2.6 A Consistency Check 34
2.7 Algebra of γ Matrices 35
2.8 An Illustrative Example 37
Review Questions 38
Problems 38
Solutions to Problems 39

3 The Propagation Kernel and Feynman Diagrams 44
3.1 The Propagation Kernel . . . 44
3.1.1 In non-relativistic Schrödinger theory . . . 44
3.1.2 In relativistic Dirac theory . . . 47
3.1.3 In momentum representation . . . 49
3.1.4 Interaction with electromagnetic field . . . 50
3.2 The Transition Amplitude . . . 51
3.2.1 First order matrix element . . . 51
3.2.2 Second order matrix element . . . 52
3.3 Feynman Diagrams . . . 53
3.3.1 In space-time representation . . . 53
3.3.2 In momentum-energy representation . . . 55
3.3.3 Basic vertices in QED . . . 56
Review Questions . . . 58
Problems . . . 58
Solutions to Problems . . . 58

4 Quantum Electrodynamics 60
4.1 Rutherford Scattering . . . 60
4.1.1 Fermi's golden rule . . . 64
4.2 Compton Scattering . . . 65
4.2.1 Kinematics . . . 66
4.2.2 Transition matrix element . . . 67
4.2.3 Transition probability . . . 70
4.2.4 Density of final states . . . 72
4.2.5 Klein-Nishina formula . . . 72
4.3 Electron-electron Scattering . . . 74
4.3.1 Evaluation of traces . . . 76
4.3.2 In the centre of momentum frame . . . 78
4.4 Electron-positron Scattering . . . 80
4.5 Electron-positron Pair Annihilation into two Photons . . . 83
4.6 Bremsstrahlung . . . 87
4.6.1 In the soft photon limit . . . 90
4.7 Electron-positron Pair Production . . . 91
4.8 Muon Pair Production in Electron-positron Collision . . . 92
4.8.1 Application to hadron production . . . 96
Review Questions . . . 97
Problems . . . 98
Solutions to Problems . . . 99

5 Radiative Corrections **109**
5.1 Electron Self Energy 109
5.2 Higher Order Corrections 114
Review Questions 117
Problems 118
Solutions to Problems 119

6 Elements of Quantum Field Theory **123**
6.1 A Brief Review of Classical Mechanics 124
6.1.1 The Lagrangian-Hamiltonian formalism 124
6.1.2 The classical fields 126
6.2 Quantization of the Field 129
6.2.1 The Schrödinger field 130
6.2.2 Quantization into Bosons 132
6.2.3 Quantization into Fermions 133
6.3 Relativistic Fields 135
Review Questions 137
Problems 138
Solutions to Problems 138

7 The Scalar Fields **142**
7.1 One-component Real Field 142
7.1.1 Fourier decomposition of the field 143
7.1.2 Quantization of the scalar field 146
7.1.3 Ground state and normal ordering 147
7.2 Complex Scalar Field 148
7.2.1 Charge-current density 150
7.2.2 Particles and antiparticles 152
7.3 The Covariant Commutation Relations 153
Review Questions 156
Problems 156
Solutions to Problems 157

8 The Dirac Field **163**
8.1 Plane Wave Solutions of the Dirac Equation 164
8.2 Lagrangian Density for the Dirac Field 166
8.3 Fourier Decomposition of the Dirac Field 167
8.4 Quantization of the Dirac Field 169
8.4.1 Anticommutator between Dirac field functions . . 171
8.5 Covariant Anticommutation Relations 172

Review Questions 175
Problems 176
Solutions to Problems 176

9 The Electromagnetic Field **179**
9.1 Maxwell's Equations 180
9.2 Electromagnetic Field Tensor 183
9.3 Quantization of the Electromagnetic Field 186
9.4 The Gupta-Bleuler Formulation 188
9.4.1 The indefinite metric operator η 190
9.4.2 The Lorentz condition 191
Review Questions 193
Problems 193
Solutions to Problems 194

10 Interaction between Fields **197**
10.1 The Tomonaga-Schwinger Equation 198
10.2 The Invariant Perturbation Theory 199
10.3 The S-matrix 202
10.4 Reduction of the S-matrix 203
10.4.1 Decomposition into normal products 203
10.4.2 Wick's chronological product 205
10.4.3 Wick's contractions 206
10.4.4 Wick's theorem 208
10.5 From S-matrix Expansion to Feynman Diagrams 209
10.5.1 Interaction with external electromagnetic field . . 209
10.5.2 Second order term in S-matrix 213
Review Questions 215
Problems 216
Solutions to Problems 216

11 Recent Developments **219**
11.1 The Renormalization Program 219
11.1.1 Regularization and Renormalization 220
11.2 Gauge Theories 221
11.2.1 Gauge theory of electromagnetic interaction 221
11.2.2 Spontaneous symmetry breaking 224
11.3 The Standard Model 228
11.3.1 Basic elementary particles - Fermions 228
11.3.2 Force mediating particles - Bosons 230

11.3.3 The Higgs Boson 230
11.3.4 Theoretical framework 231
11.3.5 Successes and failures of the Standard Model . . . 232
11.4 Gravitational Interaction 233
11.5 Birth of String Theory 234
11.5.1 Energy and Length Scales 234
11.5.2 Basic theory of strings 235
11.5.3 Open strings 236
11.5.4 Closed strings 238
11.6 Superstring Theories 238

Appendix A
List of Symbols and Notation 243

Appendix B
The Density of Final States 248

Bibliography **251**

Subject Index **253**

Chapter 1

Introduction

Classical Mechanics envisage the dichotomy of the universe – matter and waves, mass and energy and space and time. Quantum Mechanics unifies matter and waves whereas Einstein's special theory of relativity unifies mass and energy and also space and time. Relativistic Quantum Mechanics is formulated synthesizing Quantum Mechanics with the special theory of relativity.

Newton's laws of motion and Newtonian mechanics have laid the foundations for classical mechanics but it is found that the Hamiltonian and the Lagrangian formulation of classical mechanics and the Hamilton principle of least action are more suitable for treating complicated problems with constraints. The Hamiltonian theory of classical mechanics has lead to the formulation of Quantum Mechanics by treating the canonically conjugate dynamical variables, position (x, y, z) and momentum (p_x, p_y, p_z) as operators, obeying certain commutation relations.

$$[x, p_x] = i\hbar, \qquad [y, p_y] = i\hbar, \qquad [z, p_z] = i\hbar. \tag{1.1}$$

Similarly, energy E and time t are conjugate variables, obeying the commutation relation

$$[E, t] = i\hbar. \tag{1.2}$$

The commutation relations can be viewed as a consequence of Heisenberg's uncertainty relation that lead to the discrete energy levels of a bound system (*i.e.* the quantization of energy) and the quantization of angular momentum. This is what is known as *the first quantization*. The non-relativistic relation between the energy and momentum of a particle leads to the Schrödinger equation whereas the relativistic relation leads to the Klein-Gordon and Dirac equations.

1.1 The Schrödinger Equation

The non-relativistic relation for the energy E of a particle of mass m and momentum $\boldsymbol{p}$ moving in a potential V is given by

$$\frac{p^2}{2m} + V = E. \tag{1.3}$$

Treating $\boldsymbol{p}$ and E as operators

$$\boldsymbol{p} \longrightarrow -i\hbar\boldsymbol{\nabla} \qquad \text{and} \qquad E \longrightarrow i\hbar\frac{\partial}{\partial t},$$

we obtain the Schrödinger wave equation

$$\left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V\right)\Psi(\boldsymbol{r},t) = i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t), \tag{1.4}$$

where $\Psi(\boldsymbol{r},t)$ is the wave function which is interpreted as the amplitude of probability of finding the particle at position $\boldsymbol{r}$ at time t. The complex conjugate of Eq. (1.4) is

$$\left(-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V\right)\Psi^*(\boldsymbol{r},t) = -i\hbar\frac{\partial}{\partial t}\Psi^*(\boldsymbol{r},t). \tag{1.5}$$

Multiplying Eq. (1.4) by Ψ^* on the left and Eq. (1.5) by Ψ on the right and subtracting the latter from the former, we get

$$\begin{aligned} -\frac{\hbar^2}{2m}\left\{\Psi^*\boldsymbol{\nabla}^2\Psi - (\boldsymbol{\nabla}^2\Psi^*)\Psi\right\} &= i\hbar\left(\Psi^*\frac{\partial\Psi}{\partial t} + \frac{\partial\Psi^*}{\partial t}\Psi\right); \\ -\frac{\hbar^2}{2m}\boldsymbol{\nabla}\cdot\left\{\Psi^*\boldsymbol{\nabla}\Psi - (\boldsymbol{\nabla}\Psi^*)\Psi\right\} &= i\hbar\frac{\partial}{\partial t}(\Psi^*\Psi). \end{aligned} \tag{1.6}$$

Defining the probability current density $\boldsymbol{J}$ and the probability density ρ,

$$\boldsymbol{J} = \frac{\hbar}{2mi}\left\{\Psi^*\boldsymbol{\nabla}\Psi - (\boldsymbol{\nabla}\Psi^*)\Psi\right\} \qquad \text{and} \qquad \rho = \Psi^*\Psi, \tag{1.7}$$

we obtain the continuity equation

$$\frac{\partial\rho}{\partial t} + \boldsymbol{\nabla}\cdot\boldsymbol{J} = 0. \tag{1.8}$$

Equation (1.8) is the continuity equation, expressing the conservation law that the rate of change of particle density in a given region is equivalent to

the particle flux through the surface enclosing the region. Please note that the probability density ρ is a positive definite quantity which is physically acceptable.

It is remarkable that the Schrödinger equation can be used to solve a wide range of problems including bound states and scattering states[1]. Only we need to put the appropriate boundary conditions and the appropriate potentials.

1.2 The Klein-Gordon Equation

Using the relativistic energy-momentum relation for a free particle with rest mass m,

$$E^2 = p^2c^2 + m^2c^4, \tag{1.9}$$

and using the usual recipe of treating the energy and momentum as differential operators

$$\boldsymbol{p} \longrightarrow -i\hbar\boldsymbol{\nabla} \qquad \text{and} \qquad E \longrightarrow i\hbar\frac{\partial}{\partial t},$$

we obtain the relativistic wave equation

$$-\hbar^2\frac{\partial^2\Psi}{\partial t^2} = (-c^2\hbar^2\boldsymbol{\nabla}^2 + m^2c^4)\Psi, \tag{1.10}$$

which, on rearrangement, yields

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + \frac{m^2c^2}{\hbar^2}\right)\Psi = 0. \tag{1.11}$$

Equation (1.11) is *the Klein-Gordon equation*, the complex conjugate of which is given by

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2 + \frac{m^2c^2}{\hbar^2}\right)\Psi^* = 0. \tag{1.12}$$

Multiplying Eq. (1.11) by Ψ^* on the left and Eq. (1.12) by Ψ on the left and subtracting, we obtain

$$\frac{1}{c^2}\left(\Psi^*\frac{\partial^2\Psi}{\partial t^2} - \Psi\frac{\partial^2\Psi^*}{\partial t^2}\right) - \Psi^*\boldsymbol{\nabla}^2\Psi + \Psi\boldsymbol{\nabla}^2\Psi^* = 0. \tag{1.13}$$

[1]For further details, please refer to V. Devanathan, *Quantum Mechanics*, Narosa Publishing House, New Delhi (2005).

Since

$$\frac{\partial}{\partial t}\left(\Psi^*\frac{\partial\Psi}{\partial t}-\Psi\frac{\partial\Psi^*}{\partial t}\right)=\Psi^*\frac{\partial^2\Psi}{\partial t^2}-\Psi\frac{\partial^2\Psi^*}{\partial t^2}, \tag{1.14}$$

and

$$\boldsymbol{\nabla}\cdot(\Psi^*\boldsymbol{\nabla}\Psi-\Psi\boldsymbol{\nabla}\Psi^*)=\Psi^*\boldsymbol{\nabla}^2\Psi-\Psi\boldsymbol{\nabla}^2\Psi^*, \tag{1.15}$$

Eq. (1.13) can be written in the form of continuity equation

$$\frac{\partial\rho}{\partial t}+\boldsymbol{\nabla}\cdot\boldsymbol{J}=0, \tag{1.16}$$

choosing for the current density $\boldsymbol{J}$ the same expression

$$\boldsymbol{J}=\frac{\hbar}{2mi}\{\Psi^*\boldsymbol{\nabla}\Psi-(\boldsymbol{\nabla}\Psi^*)\Psi\} \tag{1.17}$$

as given in Eq. (1.7) for the non-relativistic Schrödinger equation. In the present case, the continuity equation will be satisfied only if we choose for the probability density ρ the expression

$$\rho=\frac{i\hbar}{2mc^2}\left(\Psi^*\frac{\partial\Psi}{\partial t}-\Psi\frac{\partial\Psi^*}{\partial t}\right). \tag{1.18}$$

The probability density ρ as given by Eq. (1.18) involves both Ψ and $\partial\Psi/\partial t$ which can be fixed arbitrarily and hence admits both positive and negative values. Since the probability density should be a positive-definite quantity, the Klein-Gordon equation was not accepted as a wave equation for several years until Pauli and Weisskopf[2] reinterpreted it as a field equation in the same sense as Maxwell's equation for electromagnetic field. By putting the rest mass $m=0$ in Eq. (1.11), we obtain the field equation for the electromagnetic field.

The Klein-Gordon equation is a second-order differential equation in t and this has yielded the physically unacceptable negative values also for the probability density ρ. It may be observed that the Schrödinger equation is a first-order differential equation in t and hence yielded a positive-definite value for the probability density. Taking this clue, Dirac attempted to linearize the relativistic relation $E^2=p^2c^2+m^2c^4$ which is quadratic in both E and $\boldsymbol{p}$ and arrived at the Dirac equation.

[2]W. Pauli and V. Weisskopf, *Helv. Phys. Acta.* **7**, 709 (1934).

Thus the attempts to overcome the early difficulties encountered in the formulation of Relativistic Quantum Mechanics paid rich dividends. Dirac succeeded in linearizing the relativistic relation (1.9) which is quadratic in both energy and momentum and obtained the Dirac equation for the electron with the intrinsic properties of spin and magnetic moment. Interpretation of the Klein-Gordon equation as a field equation has sowed the seed for the development of the Quantum Field theory.

1.3 The Dirac Equation

Starting with the relativistic relation $E^2 = p^2c^2 + m^2c^4$ between the energy and momentum of a free particle of mass m, Dirac obtained a linear relation

$$E = c\,\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2, \tag{1.19}$$

using two operators $\boldsymbol{\alpha}$ and β which commute with both position and momentum vectors. Squaring (1.19), we obtain

$$E^2 = c^2(\boldsymbol{\alpha}\cdot\boldsymbol{p})^2 + \beta^2 m^2c^4 + (\boldsymbol{\alpha}\cdot\boldsymbol{p})\beta mc^3 + \beta(\boldsymbol{\alpha}\cdot\boldsymbol{p})mc^3, \tag{1.20}$$

which will reduce to the relativistic energy-momentum relation $E^2 = p^2c^2 + m^2c^4$ if the operators $\boldsymbol{\alpha}$ and β obey the following relations:

$$(\boldsymbol{\alpha}\cdot\boldsymbol{p})^2 = p^2, \qquad \beta^2 = 1, \qquad \boldsymbol{\alpha}\beta = -\beta\boldsymbol{\alpha}. \tag{1.21}$$

The relation

$$(\boldsymbol{\alpha}\cdot\boldsymbol{p})^2 = (\alpha_x p_x + \alpha_y p_y + \alpha_z p_z)^2 = p^2,$$

implies

$$\begin{aligned} &\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = 1; \\ &\alpha_x\alpha_y = -\alpha_y\alpha_x; \quad \alpha_y\alpha_z = -\alpha_z\alpha_y; \quad \alpha_z\alpha_x = -\alpha_x\alpha_z. \end{aligned} \tag{1.22}$$

Using the usual recipe for the first quantization of replacing the energy and momentum by differential operators

$$E \longrightarrow i\hbar\frac{\partial}{\partial t} \qquad \text{and} \qquad \boldsymbol{p} \longrightarrow -i\hbar\boldsymbol{\nabla}$$

in Eq. (1.19), the time-dependent Dirac equation for a free particle is obtained.

$$i\hbar\frac{\partial\Psi(\boldsymbol{r},t)}{\partial t} = \left(-ic\hbar\boldsymbol{\alpha}\cdot\boldsymbol{\nabla} + \beta mc^2\right)\Psi(\boldsymbol{r},t). \tag{1.23}$$

The space-time dependence of the free particle Dirac wave function can be explicitly written as

$$\Psi(\boldsymbol{r},t) = \psi\, e^{(i/\hbar)(\boldsymbol{p}\cdot\boldsymbol{r}-Et)}. \tag{1.24}$$

This satisfies the time-dependent Dirac equation (1.23) for a free particle; from which we obtain the time-independent Dirac equation

$$(c\,\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2)\psi = E\psi, \tag{1.25}$$

with a multi-component wave function ψ. If $\alpha_x, \alpha_y, \alpha_z$ and β are matrices of dimension $N \times N$, then the Dirac wave function ψ should be a column vector with N components. In Eq. (1.25), $\boldsymbol{p}$ is the momentum vector and not an operator.

Dirac's α and β matrices

The Dirac Hamiltonian should be Hermitian, i.e., $H = H^\dagger$. Since $\boldsymbol{p}$ is Hermitian,

$$H = c(\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc); \qquad H^\dagger = c(\boldsymbol{\alpha}^\dagger\cdot\boldsymbol{p} + \beta^\dagger mc).$$

This means that $\alpha_i,\ (i = x, y, z)$ and β matrices are Hermitian.

$$\alpha_i^\dagger = \alpha_i, \quad (i = x, y, z); \qquad \beta^\dagger = \beta.$$

Since

$$\alpha_i^2 = 1, \quad (i = x, y, z) \ \text{ and } \ \beta^2 = 1,$$

it follows that

$$\alpha_i = \alpha_i^{-1} \ \text{ and } \ \beta = \beta^{-1},$$

which means that the matrices α_i and β are non-singular and consequently their determinant is non-zero.

$$\det(\alpha_i) \neq 0,\ (i = x, y, z); \qquad \det\beta \neq 0.$$

We have already seen that the Dirac matrices $\alpha_x, \alpha_y, \alpha_z, \beta$ anticommute. If they are of dimension $N \times N$, then

$$\begin{aligned} \alpha_i\beta &= -\beta\alpha_i, \\ \det(\alpha_i\beta) &= \det(-1)\det(\beta\alpha_i), \\ \det(\alpha_i)\det(\beta) &= \det(-1)\det(\beta)\det(\alpha_i), \end{aligned} \tag{1.26}$$

from which, we deduce that

$$\det(-1) \;=\; (-1)^N = 1, \tag{1.27}$$

assuming that the matrices α_i, β are of dimensions $N \times N$ and the condition (1.27) is satisfied if N is even.

$$N = 2, 4, 6, \cdots .$$

We have seen that the (2×2) matrices permit only three anticommuting matrices (the Pauli matrices) along with the unit matrix as independent matrices and all other matrices can be represented as a linear combination of them. Since in Dirac equation, we have four anticommuting matrices, they should be of higher dimension, at least (4×4).

Dirac's α and β matrices are traceless. Since $\alpha_i\beta = -\beta\alpha_i$, it follows that

$$\begin{aligned} \alpha_i\beta\alpha_i^{-1} &= -\beta \\ \text{Trace } (\alpha_i\beta\alpha_i^{-1}) &= -\text{Trace } \beta \\ \text{Trace } (\alpha_i^{-1}\alpha_i\beta) &= -\text{Trace } \beta \\ \text{Trace } (\beta) &= -\text{Trace } \beta \end{aligned} \tag{1.28}$$

This means that Trace $\beta = 0$. In the above derivation, we have used the cyclic property of the Trace.

$$\text{Trace } (ABC) = \text{ Trace } (CAB).$$

Similarly, it can be shown that Trace $\alpha_i = 0$, $(i = x, y, z)$. Let us now give an explicit representation[3] for the Dirac matrices which satisfy all the above properties.

$$\alpha_i = \begin{bmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{bmatrix}, \qquad \beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \tag{1.29}$$

where σ_i denotes the familiar Pauli matrices and I denotes the 2×2 unit matrix. Although the Dirac matrices are of dimension 4×4, we have

[3]It is the most commonly used representation but it is not the unique representation. Other representations are possible.

written conveniently in the 2×2 form in Eq. (1.29). Explicitly,

$$\alpha_x = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \qquad \alpha_y = \begin{bmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{bmatrix},$$
$$\alpha_z = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, \quad \beta = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{1.30}$$

The continuity equation

Multiplying the Dirac equation (1.23) on the left by $\Psi^\dagger$, we get

$$i\hbar\Psi^\dagger \frac{\partial \Psi}{\partial t} = -ic\hbar\Psi^\dagger \boldsymbol{\alpha} \cdot \boldsymbol{\nabla}\Psi + \beta mc^2 \Psi^\dagger \Psi. \tag{1.31}$$

Taking the Hermitian conjugate of the Dirac equation (1.23) and multiplying on the right by Ψ, we get

$$-i\hbar \frac{\partial \Psi^\dagger}{\partial t}\Psi = ic\hbar(\boldsymbol{\nabla}\Psi^\dagger) \cdot \boldsymbol{\alpha}^\dagger \Psi + \beta^\dagger mc^2 \Psi^\dagger \Psi. \tag{1.32}$$

Remembering that α and β are Hermitian matrices and subtracting (1.32) from (1.31), we obtain

$$\begin{aligned} i\hbar \frac{\partial}{\partial t}(\Psi^\dagger \Psi) &= -ic\hbar \left(\Psi^\dagger \boldsymbol{\alpha} \cdot \boldsymbol{\nabla}\Psi + (\boldsymbol{\nabla}\Psi^\dagger) \cdot \boldsymbol{\alpha}\Psi \right) \\ &= -ic\hbar \boldsymbol{\nabla} \cdot (\Psi^\dagger \boldsymbol{\alpha} \Psi), \end{aligned} \tag{1.33}$$

which can be written as a continuity equation

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0, \tag{1.34}$$

with[4]

$$\rho = \Psi^\dagger \Psi = \psi^\dagger \psi, \qquad \boldsymbol{J} = c\, \Psi^\dagger \boldsymbol{\alpha} \Psi = c\, \psi^\dagger \boldsymbol{\alpha} \psi. \tag{1.35}$$

[4]The space-time dependence of the Dirac wave function Ψ is given by Eq. (1.24) and it does not affect ρ and $\boldsymbol{J}$ as shown in Eq. (1.35).

Since ψ is a column vector and $\psi^\dagger$, a row vector as given below

$$\psi = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \qquad \psi^\dagger = \begin{bmatrix} a^* & b^* & c^* & d^* \end{bmatrix}, \tag{1.36}$$

the probability density

$$\rho = \Psi^\dagger \Psi = \psi^\dagger \psi = a^* a + b^* b + c^* c + d^* d$$

is a positive-definite quantity, overcoming the difficulty encountered in the Klein-Gordon equation. The probability current density $\boldsymbol{J}$ together with the probability density ρ obeys the continuity equation (1.34). So, the Dirac equation has become a physically acceptable relativistic wave equation.

Free particle solutions and their interpretations

The multicomponent wave functions of the Dirac equation are known as Dirac spinors and they correspond to the description of spin $\frac{1}{2}\hbar$ particle. The Dirac Hamiltonian $H = c(\boldsymbol{\alpha} \cdot \boldsymbol{p} + \beta mc)$ does not commute with the orbital angular momentum operator $\boldsymbol{L}$ and so the orbital angular momentum $\boldsymbol{L}$ is not a constant of motion but $\boldsymbol{L} + \frac{1}{2}\hbar\boldsymbol{\sigma}$ commutes with H (*vide* Problem (1.1)). Hence the total angular momentum $\boldsymbol{J} = \boldsymbol{L} + \frac{1}{2}\hbar\boldsymbol{\sigma}$ is a constant of motion, indicating that the Dirac particle has an intrinsic spin $\frac{1}{2}\hbar$. This is the greatest triumph of Dirac equation that it describes a particle such as electron with intrinsic spin $\frac{1}{2}\hbar$ and it is a consequence of the linearization of the relativistic relation $E^2 = p^2c^2 + m^2c^4$.

Although the linearization has eliminated the negative values for the probability density and thus yielded physically acceptable solutions of the Dirac equation, yet the Dirac equation permitted both positive and negative energy solutions. The negative energy solutions cannot be thrown away as unphysical since both the positive and negative energy solutions together formed a complete set.

$$\sum_{i=1}^{i=4} \psi_i \psi_i^\dagger = I. \tag{1.37}$$

The Dirac equation has four solutions, of which two of them correspond to positive energies, one with spin up and another with spin down and the

other two correspond to negative energies, one with spin up and another with spin down.

The negative energy solutions do pose a serious problem. If the negative energy states exist, then the electron in a positive energy state will make a transition to the negative energy states and it is impossible to construct any stable atom with electrons. To avoid this disastrous situation, Dirac made a bold suggestion that the negative energy states, though exist, are all filled by electrons and the Dirac vacuum[5] is one in which all the negative energy states are filled and all the positive energy states are empty. Then the Pauli exclusion principle will prevent the transition of an electron from a positive energy state to the negative energy states which are all occupied and thereby renders stability to the electron in an atom.

The hole theory of Dirac

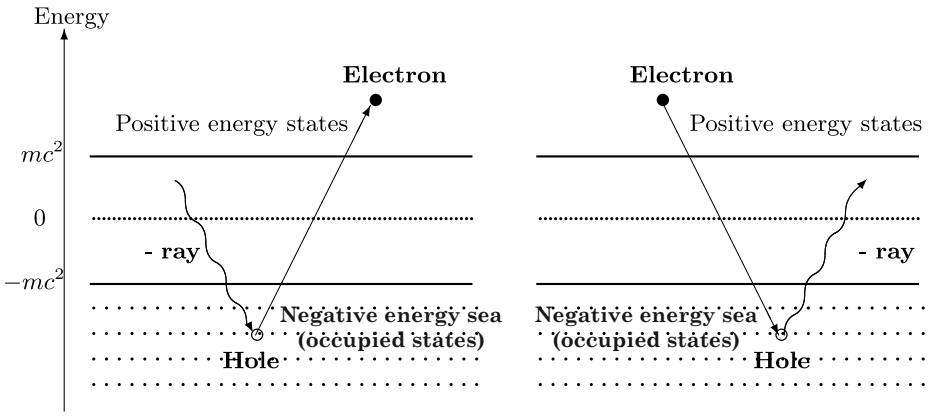
What are the consequences of this bold assumption of negative energy sea of filled electrons. A negative energy electron can absorb radiation and jump into a positive energy state; thereby creating a hole in the negative energy sea and a particle in the positive energy state. A hole in the negative energy sea is the absence of a negative energy electron from the Dirac vacuum and hence will be interpreted by the observer as the particle with opposite charge in the positive energy state. Thus the Dirac equation together with the hole theory predicts the existence of positrons which have been subsequently discovered by Anderson[6]. Thus the prediction and the subsequent discovery of positrons support the hole theory of Dirac and introduced the concept of anti-particles, which can be considered as the greatest contribution of Dirac's theory of electron.

A gamma ray of energy greater than $2mc^2$ can excite a negative energy electron to a positive energy state; thereby creating a hole and a particle as illustrated in Fig. 1.1. This process is known as *pair creation* (creation of positron and electron). The inverse process of pair annihilation (annihilation of positron and electron) accompanied by emission of gamma ray of energy greater than $2mc^2$ is also illustrated in Fig. 1.1.

Although Dirac attempted only to develop a single particle relativistic wave equation, he has ended up with the theory of many particles,

[5]Normally, we think of the vacuum as a state of emptiness but the Dirac vacuum has an infinite number of particles occupying the negative energy states!

[6]C. D. Anderson, *Phys. Rev.,* **41**, 405 (1932). C. D. Anderson was awarded the Nobel prize for physics in 1936 for the discovery of positron.



(a) Pair creation (b) Pair annihilation

Figure 1.1: Electron-positron pair creation and pair annihilation according to Dirac's hole theory

predicting pair creation and pair annihilation. Indeed the Dirac equation together with the hole theory can be truly regarded as the forerunner of *quantum field theory.*

1.4 Feynman's Positron Theory

Feynman developed an unified approach for the treatment of electrons and positrons. One can draw a world line for electron which is propagated from an initial space-time point $\boldsymbol{x}_1, t_1$ to a final space-time point $\boldsymbol{x}_2, t_2$ as shown in Fig. 1.2. It is a pictorial representation, known as the Feynman diagram. At any given time, the number of particles need not be conserved but the charge will have to be conserved. Taking this clue, Feynman intuitively identified the positron as the electron travelling backwards in time. Initially, at time t_1, there is only one electron and finally at time t_2, there is once again only one electron but at some intermediate time, t_i there are two electrons and one positron. An electron-positron pair has been created at space-time point $\boldsymbol{x}_A, t_A$, the positron being annihilated by the incoming electron at space-time point x_B, t_B and the electron emerging into the final state. Feynman describes this scenario beautifully in an inimitable way[7]. *"It is as though a bombardier flying low over a road suddenly sees three roads and it is only when two of them come*

[7]R. P. Feynman, *Phys. Rev.*, **76**, 749 (1949).

together and disappear again that he realizes that he has simply passed over a long switchback in a single road."

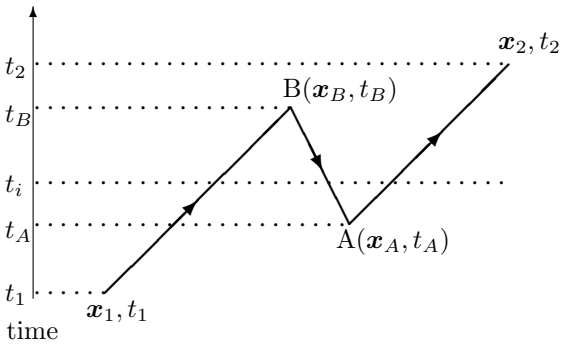


Figure 1.2: Feynman diagram representing a world line for an electron. Between the points, B and A, the electron goes back in time. It is the hole in the negative energy sea which is identified as positron that travels from A to B. At time t_1, there is only one electron and at time t_2, there is once again only one electron but at some intermediate time t_i, there are two electrons and one positron. Please note that the particle number is not conserved but the charge is conserved.

Instead of using the differential equation for the time evolution of the wave function in infinitesimal steps, Feynman developed the concept of propagation kernel which makes a giant leap forward and gives the time evolution of the wave function from one space-time point to another space-time point that is well separated. Feynman's method has greatly simplified the calculation of processes in Quantum Electrodynamics without the need of using the field-theoretical methods. In this book, we shall follow Feynman's method of solving problems in Quantum Electrodynamics and indicate how the Quantum Field Theory and interaction between fields lead to the same results through the S-matrix formalism.

1.5 Quantum Electrodynamics

In Quantum Electrodynamics (QED), we consider processes involving electrons and positrons, interacting with electromagnetic fields. Rutherford scattering, Compton scattering, electron-electron scattering, electron-positron scattering, electron-positron pair creation, electron-positron pair annihilation are some of the processes considered in this book.

In QED, the interaction between the particles and the electromagnetic fields is regarded as a perturbation. When calculations were done in the lowest order, reasonable results were obtained. Encouraged by the initial

success, attempts were made to improve the accuracy by going to higher orders but lo and behold, such attempts encountered infinities! How the infinities are tamed and handled by renormalization procedures forms an interesting anecdote and an important development.

QED was the culmination of the efforts of three theoretical physicists, Sin-Itiro Tomonaga, Julian Schwinger and Richard Feynman who shared the Nobel prize for Physics in the year 1965. Their formulations were so divergent that it was Freeman Dyson[8] who showed their equivalence. *The theory of Feynman differs profoundly in its formulation from that of Tomonaga and Schwinger. The advantages of the Feynman theory are simplicity and ease of applications, while those of Tomonaga-Schwinger are generality and theoretical completeness.*

Feynman refers to QED as *the jewel of physics - our proudest possession*, because QED is able to predict the experimental results with incredible accuracy. They are

1. The Lamb shift in hydrogen spectrum of $2S_{1/2}$ level from $2P_{1/2}$,
2. The anomalous magnetic moment of the electron,
3. Splitting of the ground state of positronium.

What I have referred to above is a success story – success in predicting all the observable quantities with great accuracy by using clever techniques to avoid the infinities. Feynman comments that *the renormalization theory is simply a way to sweep difficulties under the rug.* All the players, Tomonaga, Schwinger and Feynman feel that the theory that they have developed is intellectually not satisfactory. What they have provided is only a conservative solution but what is needed is a radical innovation and a revolutionary departure similar to what has been made in the nineteen thirties by Bohr, Heisenberg, Schrödinger and Dirac.

1.6 Quantum Field Theory

Instead of considering the Schrödinger equation, the Klein-Gordon equation and the Dirac equation as single-particle wave equations, they can be treated as classical field equations in the same way as Maxwell's equations for electromagnetic field. These field equations can be deduced using appropriate Lagrangian densities ($\mathscr{L}$) and the Euler-Lagrange equations of motion deduced from the action integral. Since the action integral is relativistically invariant, it is possible to develop a relativistically invariant theory which is manifestly covariant at every stage.

[8] F. J. Dyson, *Phys. Rev.*, **75**, 486 (1949).

Defining the conjugate momentum field $\Pi(\boldsymbol{x}, t)$

$$\Pi(\boldsymbol{x}, t) = \frac{\partial \mathscr{L}}{\partial \frac{\partial \Psi}{\partial t}}, \tag{1.38}$$

for each field $\Psi(\boldsymbol{x}, t)$, we can postulate either commutation relations or anticommutation relations for the field operator $\Psi(\boldsymbol{x}, t)$ and its conjugate momentum $\Pi(\boldsymbol{x}, t)$. The following commutation relations

$$\begin{aligned} &[\Psi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)]_- = i\hbar\delta(\boldsymbol{x}, \boldsymbol{x}'), \\ &[\Psi(\boldsymbol{x}, t), \Psi(\boldsymbol{x}', t)]_- = 0, \\ &[\Pi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)]_- = 0, \end{aligned} \tag{1.39}$$

are analogous to the Heisenberg commutation relations used in Quantum Mechanics for $\boldsymbol{x}$ and $\boldsymbol{p}$ and bear a similarity to Poisson brackets in classical mechanics. The commutation relations (1.39) are applicable to the Boson fields whereas the anticommutation relations as shown below have to be used for Fermion fields.

$$\begin{aligned} &\{\Psi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)\}_+ = i\hbar\delta(\boldsymbol{x}, \boldsymbol{x}'); \\ &\{\Psi(\boldsymbol{x}, t), \Psi(\boldsymbol{x}', t)\}_+ = 0; \\ &\{\Pi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)\}_+ = 0. \end{aligned} \tag{1.40}$$

The anticommutation relations have no classical analogue.

The commutation relations (1.39) or anticommutation relations (1.40) are used to quantize the fields and since the quantization is done for the second time, this quantization is known as *the second quantization.* It may be recalled that the Schrödinger equation, the Klein-Gordon equation and the Dirac equation are themselves obtained earlier from the corresponding Hamiltonians by treating $\boldsymbol{x}$ and $\boldsymbol{p}$ as operators obeying commutation relations. This is what we have referred to as *the first quantization.*

The field $\Psi(\boldsymbol{x}, t)$ and the conjugate momentum field $\Pi(\boldsymbol{x}, t)$ can be expanded in terms of the annihilation and creation operators (known as Fourier decomposition) and the application of commutation relations (1.39) or anticommutation relations (1.40) yields respectively the commutation relations or anticommutation relations between the annihilation and creation operators.

It is found that the Schrödinger equation, being a non-relativistic one, can be quantized by using either commutation or anticommutation relations. If commutation relations are used, then it becomes applicable to Bosons and if anticommutation relations are used then it becomes applicable to Fermions. But in the case of relativistic field equations, only one of

the relations can be used to obtain meaningful results. The Klein-Gordon equation can be quantized only be using the commutation relations between the field function and its corresponding conjugate momentum and is applicable only for Bosons. On the other hand, the Dirac equation can be quantized only by using the anticommutation relations and becomes applicable only to Fermions.

The fields interact with one another and all the phenomena that we observe are due to the interaction of fields. How the interaction of the fields can best be studied through the S-matrix formalism is outlined in Chapter 10.

The S-matrix is a perturbation series. As long as we confine ourselves to the lowest order, we get meaningful results. But, if we try to go to higher orders to improve the accuracy, we face divergent integrals. By using special techniques, known as regularization and renormalization, we can extract finite corrections. But Feynman refers to this as the shell game with the following remarks in his inimitable way.

"The shell game that we play is technically called renormalization. But no matter how clever the word, it is what I would call a dippy process! Having to resort to such hocus-pocus has prevented us from proving that the theory of quantum electrodynamics is mathematically self-consistent. I suspect that renormalization is not mathematically legitimate.... I believe there is really no satisfactory quantum electrodynamics, but I am not sure.... I think that the renormalization theory is simply a way to sweep difficulties of divergences of electrodynamics under the rug. I am, of course not sure of that!"

This scepticism was shared by many at that time. By now, it has become clearer that the renormalization techniques are legitimate methods of arriving at finite results for observable quantities and QED is indeed a renormalizable field theory. The field theory has been extended to include in its fold weak interactions and strong interactions and a Standard Model for elementary particle physics has emerged as an acceptable theory for the three of the four known elementary particle interactions. It has been the aim of theoretical physicists to formulate a unified theory which includes also gravitational interaction. It is now being felt that the Quantum Field Theory is a low-energy theory whose validity is limited to 1000 GeV, beyond which one has to look for String Theory with supersymmetry, known as Superstring Theories. This emerging scenario is discussed in the last chapter of this book.

Review Questions

1.1 What is meant by the first quantization and the second quantization? Distinguish between them and illustrate your answer by giving examples.

1.2 Obtain the Klein-Gordon equation and deduce the continuity equation that it obeys. Show that the Klein-Gordon equation permits the negative values for the probability density.

1.3 Obtain the Dirac equation and show that its solution is a multi-component wave function. Also show that the Dirac equation yields the continuity equation with positive-definite values for the probability density.

Problems

1.1 Show that the orbital angular momentum operator $\boldsymbol{L}$ does not commute with the Dirac Hamiltonian and only $\boldsymbol{L} + \frac{1}{2}\hbar\boldsymbol{\sigma}$ commutes with the Dirac Hamiltonian; thereby indicating that the Dirac equation is for a particle with intrinsic spin $\frac{1}{2}\hbar\boldsymbol{\sigma}$.

1.2 Show that the electron possesses an intrinsic magnetic moment M due to its spin and that the Dirac equation yields the correct gyromagnetic factor (g factor) $g = 2$.

1.3 Show that $\boldsymbol{\sigma}$ does not commute with the Dirac Hamiltonian but $\boldsymbol{\sigma}\cdot\boldsymbol{p}$ commutes with the Dirac Hamiltonian. Hence demonstrate that the helicity $\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|p|}$ is a good quantum number for a Dirac particle.

Solutions to Problems

1.1 It is shown that the orbital angular momentum ($\boldsymbol{L}$) is not a constant of motion since the operator $\boldsymbol{L}$ does not commute with the Dirac Hamiltonian H. However, the total angular momentum $\boldsymbol{J}$ which includes the spin $\boldsymbol{S} = \frac{1}{2}\hbar\boldsymbol{\sigma}$ is a constant of motion since $\boldsymbol{J}$ ($= \boldsymbol{L} + \boldsymbol{S}$) commutes with H.

The Dirac Hamiltonian and the orbital angular momentum operator are given by

$$H = c\,\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2, \quad \boldsymbol{L} = \boldsymbol{r}\times\boldsymbol{p}, \tag{1.41}$$

where $\boldsymbol{p}$ is to be regarded as an operator $-i\hbar\boldsymbol{\nabla}$. Since the operator $\boldsymbol{L}$ commutes with the term βmc^2 of the Dirac Hamiltonian,

$$\begin{aligned}[H, \boldsymbol{L}] &= [c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}, \boldsymbol{L}] \\ &= c[\boldsymbol{\alpha}\cdot\boldsymbol{p}, L_x]\hat{\boldsymbol{e}}_x + c[\boldsymbol{\alpha}\cdot\boldsymbol{p}, L_y]\hat{\boldsymbol{e}}_y + c[\boldsymbol{\alpha}\cdot\boldsymbol{p}, L_z]\hat{\boldsymbol{e}}_z, \end{aligned} \tag{1.42}$$

where $\hat{\boldsymbol{e}}_x, \hat{\boldsymbol{e}}_y, \hat{\boldsymbol{e}}_z$ denote the unit vectors. Let us evaluate the first term[9] on the right hand side of the above equation.

$$\begin{aligned} c[\boldsymbol{\alpha}\cdot\boldsymbol{p}, L_x] &= c[\alpha_x p_x + \alpha_y p_y + \alpha_z p_z, yp_z - zp_y] \\ &= c[\alpha_y p_y, yp_z] - c[\alpha_z p_z, zp_y] \\ &= c\alpha_y [p_y, y]p_z - c\alpha_z [p_z, z]p_y \\ &= -i\hbar c(\alpha_y p_z - \alpha_z p_y) \;=\; -i\hbar c(\boldsymbol{\alpha}\times\boldsymbol{p})_x. \end{aligned} \tag{1.43}$$

Similarly, the second and third terms can be evaluated and the final result is

$$[H, \boldsymbol{L}] = -i\hbar c(\boldsymbol{\alpha}\times\boldsymbol{p}) \neq 0. \tag{1.44}$$

Let us now evaluate the commutator bracket

$$\begin{aligned} [H, \boldsymbol{S}] &= [H, \tfrac{1}{2}\hbar\boldsymbol{\sigma}] \\ &= [c\,\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta mc^2, \tfrac{1}{2}\hbar\boldsymbol{\sigma}] \;=\; \tfrac{1}{2}\hbar c[\boldsymbol{\alpha}\cdot\boldsymbol{p}, \boldsymbol{\sigma}]. \end{aligned} \tag{1.45}$$

Using matrix algebra, we find

$$(\boldsymbol{\alpha}\cdot\boldsymbol{p})\boldsymbol{\sigma} = \begin{bmatrix} 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{bmatrix} = \begin{bmatrix} 0 & (\boldsymbol{\sigma}\cdot\boldsymbol{p})\boldsymbol{\sigma} \\ (\boldsymbol{\sigma}\cdot\boldsymbol{p})\boldsymbol{\sigma} & 0 \end{bmatrix} \tag{1.46}$$

and

$$\boldsymbol{\sigma}(\boldsymbol{\alpha}\cdot\boldsymbol{p}) = \begin{bmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{bmatrix} \begin{bmatrix} 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \end{bmatrix} = \begin{bmatrix} 0 & \boldsymbol{\sigma}(\boldsymbol{\sigma}\cdot\boldsymbol{p}) \\ \boldsymbol{\sigma}(\boldsymbol{\sigma}\cdot\boldsymbol{p}) & 0 \end{bmatrix}. \tag{1.47}$$

It can easily seen that

$$(\boldsymbol{\sigma}\cdot\boldsymbol{p})\boldsymbol{\sigma} = \boldsymbol{p} + i\boldsymbol{\sigma}\times\boldsymbol{p} \tag{1.48}$$

$$\boldsymbol{\sigma}(\boldsymbol{\sigma}\cdot\boldsymbol{p}) = \boldsymbol{p} - i\boldsymbol{\sigma}\times\boldsymbol{p} \tag{1.49}$$

and consequently

$$(\boldsymbol{\sigma}\cdot\boldsymbol{p})\boldsymbol{\sigma} - \boldsymbol{\sigma}(\boldsymbol{\sigma}\cdot\boldsymbol{p}) = 2i\boldsymbol{\sigma}\times\boldsymbol{p}. \tag{1.50}$$

Substituting the above results, the commutator bracket

$$\begin{aligned} [H, \boldsymbol{S}] &= \tfrac{1}{2}\hbar c[\boldsymbol{\alpha}\cdot\boldsymbol{p}, \boldsymbol{\sigma}] \\ &= \tfrac{1}{2}\hbar c \begin{bmatrix} 0 & 2i\boldsymbol{\sigma}\times\boldsymbol{p} \\ 2i\boldsymbol{\sigma}\times\boldsymbol{p} & 0 \end{bmatrix} \\ &= i\hbar c(\boldsymbol{\alpha}\times\boldsymbol{p}). \end{aligned} \tag{1.51}$$

[9] The matrices α_x, α_y and α_z are independent of the position coordinates and momenta and hence commute with them.

Thus, we find from Eqs. (1.44) and (1.51) that

$$[H, \boldsymbol{J}] = [H, \boldsymbol{L} + \boldsymbol{S}] = 0. \tag{1.52}$$

This means that the total angular momentum is a constant of motion for the Dirac particle and that the total angular momentum includes the spin of the Dirac particle which is $\frac{1}{2}\hbar$. *Thus the Dirac equation is for a particle of spin $\frac{1}{2}\hbar$ and it is a consequence of the linearization of the relativistic energy-momentum relation and writing it as a first order differential equation.*

1.2 The Dirac equation for a point particle of charge e in an electromagnetic field, described by the vector and scalar potentials $\boldsymbol{A}$ and ϕ, is written using the usual recipe $\boldsymbol{p} \to \boldsymbol{p} - e\boldsymbol{A}/c$, $E \to E - e\phi$.

$$\left\{ c\,\boldsymbol{\alpha} \cdot \left(\boldsymbol{p} - \frac{e\boldsymbol{A}}{c} \right) + \beta mc^2 \right\} \psi = (E - e\phi)\psi. \tag{1.53}$$

Writing ψ in the two-component form $\begin{bmatrix} \varphi_a \\ \varphi_b \end{bmatrix}$ and the matrices $\boldsymbol{\alpha}$ and β as 2×2 matrix, we obtained coupled equations in φ_a and φ_b.

$$\left\{ c \begin{bmatrix} 0 & \boldsymbol{\sigma} \cdot \boldsymbol{P} \\ \boldsymbol{\sigma} \cdot \boldsymbol{P} & 0 \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} mc^2 \right\} \begin{bmatrix} \varphi_a \\ \varphi_b \end{bmatrix} = (E - e\phi) \begin{bmatrix} \varphi_a \\ \varphi_b \end{bmatrix}, \tag{1.54}$$

using the notation $\boldsymbol{P} = \boldsymbol{p} - e\boldsymbol{A}/c$. The two coupled equations are

$$c(\boldsymbol{\sigma} \cdot \boldsymbol{P})\varphi_b + mc^2\varphi_a = (E - e\phi)\varphi_a, \tag{1.55}$$

$$c(\boldsymbol{\sigma} \cdot \boldsymbol{P})\varphi_a - mc^2\varphi_a = (E - e\phi)\varphi_a. \tag{1.56}$$

We are interested in the positive energy states with φ_a as large component and φ_b as small component. If T is the kinetic energy of the electron, then

$$E = mc^2 + T,$$

where m is the rest mass of the electron. For small T, we can go to the non-relativistic limit and investigate the Dirac equation in the presence of electromagnetic field. From Eq. (1.56), we get

$$\varphi_b = \frac{c\,\boldsymbol{\sigma} \cdot \boldsymbol{P}}{E + mc^2 - e\phi}\varphi_a \approx \frac{c\,\boldsymbol{\sigma} \cdot \boldsymbol{P}}{2mc^2}\varphi_a, \quad \text{if} \quad T,\, e\phi \ll 2mc^2. \tag{1.57}$$

Substituting this expression for φ_b in Eq. (1.55), we get

$$\frac{c^2\,(\boldsymbol{\sigma} \cdot \boldsymbol{P})^2}{2mc^2}\varphi_a = (E - mc^2 - e\phi)\varphi_a = (T - e\phi)\varphi_a. \tag{1.58}$$

The above equation is the non-relativistic limit of Dirac equation and it is an equation in φ_a. We note that[10]

$$(\boldsymbol{\sigma}\cdot\boldsymbol{P})^2 = \boldsymbol{P}\cdot\boldsymbol{P} + i\boldsymbol{\sigma}\cdot(\boldsymbol{P}\times\boldsymbol{P}), \tag{1.59}$$

$$\begin{aligned}\boldsymbol{P}\times\boldsymbol{P} &= \left(\boldsymbol{p}-\frac{e\boldsymbol{A}}{c}\right)\times\left(\boldsymbol{p}-\frac{e\boldsymbol{A}}{c}\right)\\ &= -\frac{e}{c}(\boldsymbol{A}\times\boldsymbol{p}+\boldsymbol{p}\times\boldsymbol{A}) = i\hbar\frac{e}{c}\boldsymbol{\nabla}\times\boldsymbol{A}.\end{aligned} \tag{1.60}$$

Substituting (1.59) and (1.60) in Eq. (1.58) and observing that the strength of the magnetic field $\boldsymbol{B}=\boldsymbol{\nabla}\times\boldsymbol{A}$, we get

$$\left\{\frac{1}{2m}\left(\boldsymbol{p}-\frac{e\boldsymbol{A}}{c}\right)^2 - \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\boldsymbol{B} + e\phi\right\}\varphi_a = \epsilon\varphi_a. \tag{1.61}$$

One can recognize the second term on the left hand side of the above equation as arising from the magnetic dipole moment

$$\boldsymbol{M} = \frac{e\hbar}{2mc}\boldsymbol{\sigma} = \mu_B\boldsymbol{\sigma}, \tag{1.62}$$

where $\mu_B = e\hbar/2mc$ is known as the *Bohr magneton*. Since, for the electron, e is negative, the magnetic moment is antiparallel to the spin. Moreover, since the electron spin is $\boldsymbol{S}=\frac{1}{2}\hbar\boldsymbol{\sigma}$, the magnetic moment is

$$\boldsymbol{M} = \frac{e}{mc}\boldsymbol{S}. \tag{1.63}$$

Thus, we obtain the remarkable result that the gyromagnetic ratio of the electron is e/mc or $g(e/2mc)$ with $g=2$. The correct g factor follows directly from the Dirac equation and is not put in by hand from experimental evidence.

1.3 Let us first evaluate the commutator $[\boldsymbol{\sigma}, H]_-$.

$$\begin{aligned}[\boldsymbol{\sigma}, H]_- &= [\boldsymbol{\sigma}, c(\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta\mu)]_-\\ &= c[\boldsymbol{\sigma}, \boldsymbol{\alpha}\cdot\boldsymbol{p}]_-, \quad \text{since } \boldsymbol{\sigma} \text{ commutes with } \beta\mu\\ &= -2ic(\boldsymbol{\alpha}\times\boldsymbol{p}),\end{aligned}$$

using the results (1.46) – (1.50).

[10] In deducing Eq. (1.60), problem (11.6) of my book on Quantum Mechanics will be helpful.

In the same way, the commutator $[\boldsymbol{\sigma}\cdot\boldsymbol{p}, H]_-$ can be evaluated.

$$\begin{aligned}
[\boldsymbol{\sigma}\cdot\boldsymbol{p}, H]_- &= [\boldsymbol{\sigma}\cdot\boldsymbol{p}, c(\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta\mu)]_- \\
&= c[\boldsymbol{\sigma}\cdot\boldsymbol{p}, \boldsymbol{\alpha}\cdot\boldsymbol{p}]_- \\
&= c\left\{\begin{bmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p}\end{bmatrix}\begin{bmatrix} 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ -\boldsymbol{\sigma}\cdot\boldsymbol{p} & 0\end{bmatrix}\right. \\
&\quad \left. -\begin{bmatrix} 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ -\boldsymbol{\sigma}\cdot\boldsymbol{p} & 0\end{bmatrix}\begin{bmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p}\end{bmatrix}\right\} \\
&= c\left\{\begin{bmatrix} 0 & (\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 \\ -(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 & 0\end{bmatrix} - \begin{bmatrix} 0 & (\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 \\ -(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 & 0\end{bmatrix}\right\} \\
&= 0.
\end{aligned}$$

Since $\boldsymbol{\sigma}\cdot\boldsymbol{p}$ commutes with the Dirac Hamiltonian H, one can find simultaneously the eigenvalues of both. Since

$$\left(\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|p|}\right)^2 = 1,$$

the eigenvalue of the operator $\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|p|}$, known as the Helicity operator is

$$\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{|p|} = \pm 1.$$

Chapter 2

The Dirac Equation and its Solutions

There are two ways of writing the Dirac equation. One is the conventional way of writing as given by Eq. (2.2), using the Dirac matrices α and β and the other is the Feynman way of writing as given by Eq. (2.37), using the γ matrices . They differ by a multiplicative factor $\beta(\equiv \gamma_0)$. Their positive energy solutions coincide but their negative energy solutions differ by a change in sign of the momentum vector $\boldsymbol{p}$. In other words, the negative energy solutions of the conventional Dirac equation correspond to electron of momentum $\boldsymbol{p}$ and energy $-E$ whereas the negative energy solutions of Dirac equation in Feynman notation correspond to electron of momentum $-\boldsymbol{p}$ and energy $-E$. The reversal in the sign of momentum for the negative energy electron has been interpreted by Feynman as the reversal of time. Feynman made a bold suggestion that the negative energy electron travels backward in time which is equivalent to the positive energy positron travelling forward in time. This new viewpoint has considerably simplified the calculation of electrodynamic processes which can be elegantly represented by means of Feynman diagrams.

2.1 The free particle solutions

Starting with the Dirac Hamiltonian of a free particle with momentum $\boldsymbol{p}$ and rest mass m

$$H = c\boldsymbol{\alpha} \cdot \boldsymbol{p} + \beta mc^2, \tag{2.1}$$

which is linear in energy and momentum, we can write down the time-independent Dirac equation as

$$(c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2)\psi = E\psi. \tag{2.2}$$

Although $\boldsymbol{\alpha}$ and β are 4×4 matrices as given in (1.30) (Eq.1.29), it is convenient to write them in the compressed form as 2×2 matrices as shown in Eq. (1.29) (Eq. 1.28) for easy algebraic manipulation.

Writing the Dirac equation (2.2) in a more simplified form $A\psi = 0$, a non-trivial solution for ψ can be obtained by imposing the condition, $\det A = 0$.

$$\begin{aligned} \det A &= \det\left(c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2 - EI\right) \\ &= \begin{vmatrix} mc^2-E & c\,\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ c\,\boldsymbol{\sigma}\cdot\boldsymbol{p} & -(mc^2+E) \end{vmatrix} = 0, \end{aligned} \tag{2.3}$$

where I is the unit matrix and $\boldsymbol{\sigma}\,(\sigma_x,\sigma_y,\sigma_z)$ are Pauli matrices. This leads to the relativistic relation $E^2 = p^2c^2+m^2c^4$ and guarantees that the Dirac equation satisfies the relativistic relation. So, the Dirac equation is only a linearized form of the relativistic energy–momentum relation in operator formalism. It follows that the Dirac Hamiltonian has two eigenvalues, one positive and the other negative.

$$E = \pm\sqrt{p^2c^2+m^2c^4} = \pm c\,W, \tag{2.4}$$

with

$$W = \sqrt{p^2+m^2c^2} = \sqrt{p^2+\mu^2},\ \text{using the symbol}\ \mu = mc. \tag{2.5}$$

In this choice, W is always a positive quantity.

Writing the Dirac equation

$$(c\,\boldsymbol{\alpha}\cdot\boldsymbol{p}+\beta mc^2)\psi = E\psi,$$

as a 2×2 matrix equation

$$\begin{bmatrix} mc^2 & c\,\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ c\,\boldsymbol{\sigma}\cdot\boldsymbol{p} & -mc^2 \end{bmatrix} \begin{bmatrix} \varphi_a \\ \varphi_b \end{bmatrix} = E \begin{bmatrix} \varphi_a \\ \varphi_b \end{bmatrix}, \tag{2.6}$$

where we have denoted the Dirac wave function in a two component form $\begin{bmatrix} \varphi_a \\ \varphi_b \end{bmatrix}$, we obtain a coupled equation in φ_a and φ_b.

$$mc^2\,\varphi_a + c\,\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\varphi_b = E\,\varphi_a. \tag{2.7}$$

$$c\,\boldsymbol{\sigma}\cdot\boldsymbol{p}\,\varphi_a - mc^2\,\varphi_b = E\,\varphi_b. \tag{2.8}$$

The energy E can assume both positive and negative values. The coupled equations (2.7) and (2.8) give only the ratios φ_a/φ_b. From Eq. (2.7), we obtain

$$\frac{\varphi_a}{\varphi_b} = \frac{c\,\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E - mc^2} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{\pm W - \mu}, \tag{2.9}$$

which becomes indeterminate as $\boldsymbol{p} \to 0$ for positive energy state since both the components φ_a and φ_b become zero. So, Eq. (2.9) is relevant only for negative energy states. In a similar way, it can be deduced from Eq. (2.8) that

$$\frac{\varphi_b}{\varphi_a} = \frac{c\,\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E + mc^2} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{\pm W + \mu} \tag{2.10}$$

becomes indeterminate for negative energy states as $\boldsymbol{p} \to 0$ and so applicable only for positive energy states.

So, we shall choose solution (2.10) for positive energy states and solution (2.9) for negative energy states. Since they give only the ratios, one is free to choose either φ_a or φ_b. The obvious choice is to choose one of them as Pauli spinor,

$$\chi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{or} \quad \chi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

For positive energy solutions, let us choose $\varphi_a = \chi_+$ or χ_- and for negative energy solutions, we shall choose $\varphi_b = \chi_+$ or χ_-. Thus, we obtain the following four solutions:

$$\psi_1 = \begin{bmatrix} \chi_+ \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{W+\mu}\chi_+ \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \frac{p_z}{W+\mu} \\ \frac{p_+}{W+\mu} \end{bmatrix}; \quad \psi_2 = \begin{bmatrix} \chi_- \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{W+\mu}\chi_- \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ \frac{p_-}{W+\mu} \\ \frac{-p_z}{W+\mu} \end{bmatrix}; \tag{2.11}$$

$$\psi_3 = \begin{bmatrix} \frac{-\boldsymbol{\sigma}\cdot\boldsymbol{p}}{W+\mu}\chi_+ \\ \chi_+ \end{bmatrix} = \begin{bmatrix} \frac{-p_z}{W+\mu} \\ \frac{-p_+}{W+\mu} \\ 1 \\ 0 \end{bmatrix}; \quad \psi_4 = \begin{bmatrix} \frac{-\boldsymbol{\sigma}\cdot\boldsymbol{p}}{W+\mu}\chi_- \\ \chi_- \end{bmatrix} = \begin{bmatrix} \frac{-p_-}{W+\mu} \\ \frac{p_z}{W+\mu} \\ 0 \\ 1 \end{bmatrix}; \tag{2.12}$$

with

$$p_+ = p_x + ip_y \quad \text{and} \quad p_- = p_x - ip_y.$$

In obtaining the above solutions, we have made use of the following relations:

$$\begin{aligned}
\boldsymbol{\sigma}\cdot\boldsymbol{p} &= \sigma_x p_x + \sigma_y p_y + \sigma_z p_z \\
&= \begin{bmatrix} p_z & p_x - ip_y \\ p_x + ip_y & -p_z \end{bmatrix} = \begin{bmatrix} p_z & p_- \\ p_+ & -p_z \end{bmatrix}. \qquad (2.13)
\end{aligned}$$

$$\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{W+\mu}\chi_+ = \frac{1}{W+\mu}\begin{bmatrix} p_z & p_- \\ p_+ & -p_z \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{p_z}{W+\mu} \\ \frac{p_+}{W+\mu} \end{bmatrix}. \qquad (2.14)$$

$$\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{W+\mu}\chi_- = \frac{1}{W+\mu}\begin{bmatrix} p_z & p_- \\ p_+ & -p_z \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{p_-}{W+\mu} \\ \frac{-p_z}{W+\mu} \end{bmatrix}. \qquad (2.15)$$

The four solutions ψ_1, ψ_2, ψ_3 and ψ_4 of the Dirac equation are called Dirac's four-spinors since the Dirac equation corresponds to spin-$1/2$ particle. The solutions ψ_1 and ψ_2 correspond to positive energy states and ψ_3 and ψ_4 correspond to negative energy states; of which ψ_1 and ψ_3 denote the spin-up state and ψ_2 and ψ_4 denote the spin-down state.

2.2 Orthogonal and Closure Properties

It can be easily seen that the four solutions are mutually orthogonal. They can also be normalized such that

$$\langle\psi_i|\psi_j\rangle = \psi_i^\dagger\psi_j = \delta_{ij}, \qquad i,j = 1,2,3,4.$$

The normalized solutions of the free particle Dirac equation are

$$\psi_1 = N\begin{bmatrix} 1 \\ 0 \\ \frac{p_z}{W+\mu} \\ \frac{p_+}{W+\mu} \end{bmatrix}; \qquad \psi_2 = N\begin{bmatrix} 0 \\ 1 \\ \frac{p_-}{W+\mu} \\ \frac{-p_z}{W+\mu} \end{bmatrix}; \qquad (2.16)$$

$$\psi_3 = N\begin{bmatrix} \frac{-p_z}{W+\mu} \\ \frac{-p_+}{W+\mu} \\ 1 \\ 0 \end{bmatrix}; \qquad \psi_4 = N\begin{bmatrix} \frac{-p_-}{W+\mu} \\ \frac{p_z}{W+\mu} \\ 0 \\ 1 \end{bmatrix}; \qquad (2.17)$$

with the normalization factor

$$N = \sqrt{\frac{W+\mu}{2W}}.$$

Let us now arrange all the four solutions of the Dirac equation for the free particle side by side to form a matrix $D(\boldsymbol{p})$.

$$D(\boldsymbol{p}) = \left(\frac{W+\mu}{2W}\right)^{1/2} \begin{bmatrix} 1 & 0 & \frac{-p_z}{W+\mu} & \frac{-p_-}{W+\mu} \\ 0 & 1 & \frac{-p_+}{W+\mu} & \frac{p_z}{W+\mu} \\ \frac{p_z}{W+\mu} & \frac{p_-}{W+\mu} & 1 & 0 \\ \frac{p_+}{W+\mu} & \frac{-p_z}{W+\mu} & 0 & 1 \end{bmatrix}. \tag{2.18}$$

If $\psi_1(0), \psi_2(0), \psi_3(0)$ and $\psi_4(0)$ are the fundamental four-spinors in the limiting case $\boldsymbol{p} \to 0$,

$$\psi_1(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \psi_2(0) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \psi_3(0) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \psi_4(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \tag{2.19}$$

then it follows that

$$\psi_i(\boldsymbol{p}) = D(\boldsymbol{p})\psi_i(0). \tag{2.20}$$

It can be easily verified that the matrix $D(\boldsymbol{p})$ can be written in a compact form as

$$D(\boldsymbol{p}) = \left(\frac{W+\mu}{2W}\right)^{1/2} \left(1 - \frac{\beta(\boldsymbol{\alpha}\cdot\boldsymbol{p})}{W+\mu}\right). \tag{2.21}$$

It is to be observed that $\boldsymbol{p}$ in $D(\boldsymbol{p})$ is not an operator. Equation (2.20) enables you to go from the free particle solutions given in the rest frame to any other frame of reference in which the particle has momentum $\boldsymbol{p}$. So, $D(\boldsymbol{p})$ is known as the Lorentz boost which tells you how the Dirac wave function transforms under Lorentz transformation.

All the four solutions ψ_1, ψ_2, ψ_3 and ψ_4 together form a complete set since

$$\sum_{i=1}^{4} |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^{4} \psi_i\psi_i^\dagger = I. \tag{2.22}$$

This is known as the closure property. That is why the negative energy solutions ψ_3 and ψ_4 cannot be discarded as unphysical and the attempt to explain them has led to *the important concept of antiparticles.*

2.3 Projection Operators

Instead of summing over all the four states in Eq. (2.22), a partial sum over positive or negative energy states has a nice physical interpretation as projection operator for positive or negative energy states.

$$\sum_{i=1,2} |\psi_i\rangle\langle\psi_i| = \psi_1\psi_1^\dagger + \psi_2\psi_2^\dagger \;=\; \Lambda^+, \tag{2.23}$$

$$\sum_{i=3,4} |\psi_i\rangle\langle\psi_i| = \psi_3\psi_3^\dagger + \psi_4\psi_4^\dagger \;=\; \Lambda^-, \tag{2.24}$$

where Λ^+ and Λ^- are known as projection operators for positive and negative energy states. Substituting the normalized state vectors (2.16) and (2.17) in Eqs. (2.23) and (2.24), we obtain expressions for the projection operators.

$$\Lambda^+ \;=\; \frac{1}{2}\left(I + \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta\mu}{W}\right). \tag{2.25}$$

$$\Lambda^- \;=\; \frac{1}{2}\left(I - \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta\mu}{W}\right). \tag{2.26}$$

It can be easily verified that Λ^+ acting on positive energy states ψ_1 or ψ_2 yields unity whereas Λ^+ acting on negative energy states ψ_3 or ψ_4 yields zero. That is why Λ^+ is called the projection operator for positive energy states. An opposite feature is exhibited by Λ^-.

$$\Lambda^+\psi_i = \frac{1}{2}\left(I + \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta\mu}{W}\right)\psi_i \;=\; \begin{cases} \psi_i, & i = 1,2; \\ 0, & i = 3,4. \end{cases} \tag{2.27}$$

$$\Lambda^-\psi_i = \frac{1}{2}\left(I - \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta\mu}{W}\right)\psi_i \;=\; \begin{cases} 0, & i = 1,2; \\ \psi_i, & i = 3,4. \end{cases} \tag{2.28}$$

It can be easily verified that

$$\Lambda^+ + \Lambda^- = I. \tag{2.29}$$

2.4 Sum over Spin States

Let us now consider the scattering of a Dirac particle from an initial state ψ_i to a final state ψ_f due to an interaction represented by the operator $\mathcal{O}$. The transition matrix element T_{fi} is given by[1]

$$T_{fi} = (\bar{\psi}_f\, \mathcal{O}\, \psi_i), \tag{2.30}$$

[1]In non-relativistic physics, we usually write $\psi^\dagger\mathcal{O}\psi$ but in relativistic physics, we should write $\bar{\psi}\mathcal{O}\psi$ since it is $\bar{\psi}\psi$ and not $\psi^\dagger\psi$, that transforms as a Lorentz scalar.

where $\bar{\psi}_f = \psi_f^\dagger \gamma_0$ and $\gamma_0 \equiv \beta$. If the spins of both the incident and scattered particles are not observed, the transition probability is given by the square of the matrix element, obtained by summing over the final spin states and averaging over the initial spin states.

$$\begin{aligned} |T_{fi}|^2 &= \frac{1}{2}\sum_{i,f}(\bar{\psi}_f \mathcal{O} \psi_i)(\bar{\psi}_f \mathcal{O} \psi_i)^\dagger \\ &= \frac{1}{2}\sum_{i,f}(\psi_f^\dagger \gamma_0 \mathcal{O} \psi_i)(\psi_i^\dagger \mathcal{O}^\dagger \gamma_0 \psi_f), \end{aligned} \tag{2.31}$$

where the summation indices i, f are over the two spin states denoted by ψ_1 and ψ_2, corresponding to positive energy states only. The factor $\frac{1}{2}$ is due to averaging over the two initial spin states ψ_1 and ψ_2. Using the algebra of matrix multiplication,

$$\begin{aligned} |T_{fi}|^2 &= \frac{1}{2}\sum_{i,f}\sum_{\rho,\lambda,\rho',\lambda'}\{(\psi_f^\dagger)_\rho\,(\gamma_0\mathcal{O})_{\rho\lambda}\,(\psi_i)_\lambda\}\{(\psi_i^\dagger)_{\rho'}\,(\mathcal{O}^\dagger\gamma_0)_{\rho'\lambda'}(\psi_f)_{\lambda'}\} \\ &= \frac{1}{2}\sum_{i,f}\sum_{\rho,\lambda,\rho',\lambda'}(\psi_f^\dagger)_\rho\,(\gamma_0\mathcal{O})_{\rho\lambda}\,(\psi_i\psi_i^\dagger)_{\lambda\rho'}\,(\mathcal{O}^\dagger\gamma_0)_{\rho'\lambda'}(\psi_f)_{\lambda'} \\ &= \frac{1}{2}\sum_{i,f}\sum_{\rho,\lambda,\rho',\lambda'}(\gamma_0\mathcal{O})_{\rho\lambda}\,(\psi_i\psi_i^\dagger)_{\lambda\rho'}\,(\mathcal{O}^\dagger\gamma_0)_{\rho'\lambda'}(\psi_f\psi_f^\dagger)_{\lambda'\rho} \\ &= \frac{1}{2}\sum_{\rho,\lambda,\rho',\lambda'}(\gamma_0\mathcal{O})_{\rho\lambda}\,(\Lambda_i^+)_{\lambda\rho'}\,(\mathcal{O}^\dagger\gamma_0)_{\rho'\lambda'}(\Lambda_f^+)_{\lambda'\rho} \\ &= \frac{1}{2}\mathrm{Tr}\,(\gamma_0\,\mathcal{O}\,\Lambda_i^+\,\mathcal{O}^\dagger\,\gamma_0\,\Lambda_f^+). \end{aligned} \tag{2.32}$$

In the above equation, Λ_i^+ and Λ_f^+ are the projection operators of the initial and final particle, as defined in Eq. (2.27), obtained after summing over the two spin states corresponding to the positive energy state.

$$\Lambda_i^+ = \frac{1}{2}\left(I + \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}_i + \beta m}{W_i}\right), \tag{2.33}$$

$$\Lambda_f^+ = \frac{1}{2}\left(I + \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p}_f + \beta m}{W_f}\right). \tag{2.34}$$

2.5 In Feynman's Notation

Multiply the Dirac equation (2.2) by β from the left

$$(c\beta\boldsymbol{\alpha}\cdot\boldsymbol{p} + mc^2)\psi = \beta E\psi. \tag{2.35}$$

Using the natural units $\hbar = 1, c = 1$ and introducing γ matrices

$$\gamma_0 = \beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \qquad \boldsymbol{\gamma} = \beta\boldsymbol{\alpha} = \begin{bmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{bmatrix}, \tag{2.36}$$

we can rewrite the Dirac equation after rearrangement in the form

$$\not{p}\,\psi = m\,\psi, \qquad \text{with} \qquad \not{p} = \gamma_0 E - \boldsymbol{\gamma}\cdot\boldsymbol{p}. \tag{2.37}$$

Writing the characteristic equation for the matrix $\not{p}$,

$$\begin{vmatrix} E-\lambda & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -E-\lambda \end{vmatrix} = 0, \tag{2.38}$$

we get

$$-(E^2 - \lambda^2) + (\boldsymbol{\sigma}\cdot\boldsymbol{p})^2 = 0,$$

from which we obtain the eigenvalues of the matrix $\not{p}$.

$$\lambda^2 = E^2 - p^2 = m^2; \qquad \lambda = \pm m. \tag{2.39}$$

From (2.39), we find that E can have both positive and negative values.

$$E^2 = p^2 + m^2; \qquad E = \pm\sqrt{p^2+m^2} = \pm E_p; \tag{2.40}$$

where E_p is always a positive quantity. The two eigenvalues of $\not{p}$ obtained in Eq. (2.39) lead to two equations, one for 'positive eigenvalue' state and the other for 'negative eigenvalue' state.

$$\not{p}\,\psi_p = m\,\psi_p, \tag{2.41}$$

$$\not{p}\,\psi_n = -m\,\psi_n, \tag{2.42}$$

where ψ_p and ψ_n denote the positive and negative eigenvalue states. The equation for ψ_n is obtained by reversing the sign of energy and momentum so that $\not{p}$ is changed into $-\not{p}$. *The state ψ_n which represents the negative energy electron with momentum $-\boldsymbol{p}$ is to be associated with the state of a positron with positive energy and momentum $+\boldsymbol{p}$, according to the Dirac hole theory.*

2.5.1 Positive eigenvalue states

Writing Eq. (2.41) in the 2×2 matrix form, choosing a two component form $\begin{bmatrix} \phi_a \\ \phi_b \end{bmatrix}$ for ψ_p, we get

$$\begin{bmatrix} E_p & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -E_p \end{bmatrix} \begin{bmatrix} \phi_a \\ \phi_b \end{bmatrix} = m \begin{bmatrix} \phi_a \\ \phi_b \end{bmatrix}, \tag{2.43}$$

which yields two coupled equations in ϕ_a and ϕ_b.

$$E_p\phi_a - \boldsymbol{\sigma}\cdot\boldsymbol{p}\phi_b = m\phi_a \quad \longrightarrow \quad \frac{\phi_a}{\phi_b} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p - m} \tag{2.44}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\phi_a - E_p\phi_b = m\phi_b \quad \longrightarrow \quad \frac{\phi_b}{\phi_a} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + m} \tag{2.45}$$

As $\boldsymbol{p} \to 0$, Eq. (2.44) becomes indeterminate. So, we choose the solution (2.45) for positive eigenvalue states with $\phi_a = \chi_\pm$. Hence

$$\psi_p = \begin{bmatrix} \chi_\pm \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_\pm \end{bmatrix}. \tag{2.46}$$

2.5.2 Negative eigenvalue states

In a similar way, let us obtain two coupled equations for ψ_n choosing a two component form $\begin{bmatrix} \phi_A \\ \phi_B \end{bmatrix}$ for ψ_n.

$$\begin{bmatrix} E_p & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -E_p \end{bmatrix} \begin{bmatrix} \phi_A \\ \phi_B \end{bmatrix} = -m \begin{bmatrix} \phi_A \\ \phi_B \end{bmatrix}, \tag{2.47}$$

The coupled equations yield the ratio ϕ_A/ϕ_B.

$$E_p\phi_A - \boldsymbol{\sigma}\cdot\boldsymbol{p}\phi_B = -m\phi_A \quad \longrightarrow \quad \frac{\phi_A}{\phi_B} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + m} \tag{2.48}$$

$$\boldsymbol{\sigma}\cdot\boldsymbol{p}\phi_A - E_p\phi_B = -m\phi_B \quad \longrightarrow \quad \frac{\phi_B}{\phi_A} = \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p - m} \tag{2.49}$$

Since (2.49) becomes indeterminate as $\boldsymbol{p} \to 0$, we shall choose the solution (2.48) with $\phi_B = \chi_\pm$. Hence

$$\psi_n = \begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_\pm \\ \chi_\pm \end{bmatrix}. \tag{2.50}$$

Earlier, we have normalized the Dirac wave functions $\psi^\dagger\psi$ to 1 but this normalization is not relativistically invariant. Since $\psi^\dagger\psi$ (which is the probability density) transforms as the zeroth component of a four-vector current, it is possible to make a relativistically invariant normalization

by setting it equal to the zeroth component of a suitable four-vector, say, energy-momentum four-vector. Feynman[2] has chosen the normalization[3]

$$\psi_p^\dagger \psi_p = 2E_p \text{ or equivalently } \bar{\psi}_p \psi_p = \psi_p^\dagger \gamma_0 \psi_p = 2m, \tag{2.51}$$

for positive eigenvalue solutions. The normalized solutions ψ_p are

$$\psi_p = \sqrt{E_p + m} \begin{bmatrix} \chi_\pm \\ \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E_p + m} \chi_\pm \end{bmatrix}. \tag{2.52}$$

It can be easily verified that

$$\sum_{\text{spins}} \psi_p \tilde{\psi}_p = \not{p} + m. \tag{2.53}$$

The normalized negative eigenvalue solutions are

$$\psi_n = \sqrt{E_p + m} \begin{bmatrix} \frac{\boldsymbol{\sigma} \cdot \boldsymbol{p}}{E_p + m} \chi_\pm \\ \chi_\pm \end{bmatrix}, \tag{2.54}$$

the normalization being

$$\psi_n^\dagger \psi_n = 2E_p; \text{ or equivalently } \bar{\psi}_n \psi_n = \psi_n^\dagger \gamma_0 \psi_n = -2m. \tag{2.55}$$

It can be easily verified by matrix multiplication that

$$\sum_{\text{spins}} \psi_n \bar{\psi}_n = \not{p} - m. \tag{2.56}$$

It can be verified that

$$\sum_{\text{spins}} \{\psi_p \bar{\psi}_p - \psi_n \bar{\psi}_n\} = 2mI. \tag{2.57}$$

The operators $\not{p}+m$ and $\not{p}-m$ are the projection operators for the positive and negative eigenvalue states.

$$\begin{aligned} (\not{p} + m)\psi_p &= 2m\psi_p, \\ (\not{p} + m)\psi_n &= 0, \\ (\not{p} - m)\psi_p &= 0, \\ (\not{p} - m)\psi_n &= -2m\psi_n. \end{aligned} \tag{2.58}$$

[2] R. P. Feynman, *Quantum Electrodynamics*, W. A. Benjamin, Inc., New York (1962).

[3] Schweber et al. (1956) choose a slightly different normalization $\bar{\psi}_p \psi_p = 1$ and $\bar{\psi}_n \psi_n = -1$ such that

$$\sum_{\text{spins}} (\tilde{\psi}_p \psi_p - \tilde{\psi}_n \psi_n) = 4 \quad \text{and} \quad \sum_{\text{spins}} (\psi_p \tilde{\psi}_p - \psi_n \tilde{\psi}_n) = I.$$

Let us now write down explicitly the normalized Feynman wave functions for the two spin orientations[4] of the positive and negative eigenvalue states.

$$\psi_1 = \psi_{p\uparrow} = N \begin{bmatrix} 1 \\ 0 \\ \frac{p_z}{E_p+m} \\ \frac{p_+}{E_p+m} \end{bmatrix}; \qquad \psi_2 = \psi_{p\downarrow} = N \begin{bmatrix} 0 \\ 1 \\ \frac{p_-}{E_p+m} \\ \frac{-p_z}{E_p+m} \end{bmatrix}; \quad (2.59)$$

$$\psi_3 = \psi_{n\uparrow} = N \begin{bmatrix} \frac{p_z}{E_p+m} \\ \frac{p_+}{E_p+m} \\ 1 \\ 0 \end{bmatrix}; \qquad \psi_4 = \psi_{n\downarrow} = N \begin{bmatrix} \frac{p_-}{E_p+m} \\ \frac{-p_z}{E_p+m} \\ 0 \\ 1 \end{bmatrix}; \quad (2.60)$$

with the normalization factor

$$N = \sqrt{E_p + m}.$$

Let us now arrange all the four solutions of the Feynman equation for the free particle side by side to form a matrix $S(\boldsymbol{p})$.

$$S(\boldsymbol{p}) = (E_p + m)^{1/2} \begin{bmatrix} 1 & 0 & \frac{p_z}{E_p+m} & \frac{p_-}{E_p+m} \\ 0 & 1 & \frac{p_+}{E_p+m} & \frac{-p_z}{E_p+m} \\ \frac{p_z}{E_p+m} & \frac{p_-}{E_p+m} & 1 & 0 \\ \frac{p_+}{E_p+m} & \frac{-p_z}{E_p+m} & 0 & 1 \end{bmatrix}. \quad (2.61)$$

It can be easily verified that the matrix $S(\boldsymbol{p})$ can be written in a compact form as

$$S(\boldsymbol{p}) = \sqrt{E_p + m} \left(1 + \frac{\boldsymbol{\alpha} \cdot \boldsymbol{p}}{E_p + m}\right) = \sqrt{E_p + m} \left(1 + \frac{\gamma_0 \boldsymbol{\gamma} \cdot \boldsymbol{p}}{E_p + m}\right). \quad (2.62)$$

If $\psi_1(0), \psi_2(0), \psi_3(0)$ and $\psi_4(0)$ are the fundamental four-spinors as given by Eq. (2.19) in the limiting case $\boldsymbol{p} \to 0$, then

$$\psi_i(\boldsymbol{p}) = S(\boldsymbol{p})\psi_i(0). \quad (2.63)$$

Equation (2.63) enables you to go from the free particle solutions given in the rest frame to any other frame of reference in which the particle has momentum $\boldsymbol{p}$. So, $S(\boldsymbol{p})$ is the Lorentz boost operator.

[4]The two spin orientations, spin up and spin down are denoted by arrow marks $\uparrow$ and $\downarrow$ in Eqs. (2.59) and (2.60).

It may be observed that the positive energy spinors of Dirac given by Eq. (2.16) coincide with the positive eigenvalue spinors of Feynman given by Eq. (2.59), except for normalization factor; whereas the negative energy spinors of Dirac given by Eq. (2.17) differ from the negative eigenvalue spinors of Feynman given by Eq. (2.60) with respect to the sign of the momentum vector $\boldsymbol{p}$. The source of this discrepancy can easily be traced. The negative energy solutions of the Dirac equation are obtained by changing the sign of energy alone and not momentum, whereas in Feynman's negative eigenvalue equation, the signs of both energy and momentum are reversed. A hole in the negative energy sea of electrons with energy $-E_p$ and momentum $-\boldsymbol{p}$ is equivalent to a positron with energy E_p and momentum $\boldsymbol{p}$. In other words, the absence of an electron with energy $-E_p$ and momentum $-\boldsymbol{p}$ in the negative energy sea is interpreted as a positron with energy E_p and momentum $\boldsymbol{p}$. If the spin projection of the negative energy electron is $+\frac{1}{2}$, then the associated positron will have spin projection $-\frac{1}{2}$.

For the sake of clarity, we can rewrite the four solutions of the Dirac equations (2.59) and (2.60) in Feynman notation as

$$\psi_1 = \psi_{p\uparrow}(E_p, \boldsymbol{p}) = u_\uparrow(E_p, \boldsymbol{p}) = u_1(E_p, \boldsymbol{p}), \tag{2.64}$$

$$\psi_2 = \psi_{p\downarrow}(E_p, \boldsymbol{p}) = u_\downarrow(E_p, \boldsymbol{p}) = u_2(E_p, \boldsymbol{p}), \tag{2.65}$$

$$\psi_3 = \psi_{n\uparrow}(-E_p, -\boldsymbol{p}) = u_\uparrow(-E_p, -\boldsymbol{p}) = v_\downarrow(E_p, \boldsymbol{p}) = v_2(E_p, \boldsymbol{p}), \tag{2.66}$$

$$\psi_4 = \psi_{n\downarrow}(-E_p, -\boldsymbol{p}) = u\downarrow(-E_p, -\boldsymbol{p}) = v_\uparrow(E_p, \boldsymbol{p}) = v_1(E_p, \boldsymbol{p}). \tag{2.67}$$

In the above equations, we have introduced a new notation, replacing the positive energy solutions by electron spinors u_1, u_2 and the negative energy solutions by positron spinors v_1, v_2.

Generally, ψ_i, $(i = 1, 2, 3, 4)$ is used denote the plane wave solutions of the Dirac equation for free particle.

$$\psi_i = u_i e^{i\boldsymbol{p}\cdot\boldsymbol{x}} \qquad \text{or} \qquad \psi_i = u_i e^{i(\boldsymbol{p}\cdot\boldsymbol{x} - Et)} = u_i e^{-i\mathbf{p}\cdot\mathbf{x}}, \tag{2.68}$$

depending upon whether we wish to represent the plane wave in the three-dimensional space or four-dimensional space. The four dimensional space-time is preferred for describing any phenomenon in a Lorentz invariant way. Bold italics $\boldsymbol{p}$, $\boldsymbol{x}$ are used to denote the three-dimensional vectors and bold upright letters $\mathbf{p}$, $\mathbf{x}$ are used to denote the four-vectors (energy-momentum four vector and the space-time four-vector). The plane wave part is suppressed in our discussion till now and we have concentrated only on the spinor part. In the application to Quantum Electrodynamics,

we find it convenient to use the symbols u_1, u_2 to denote electron spinors and v_1, v_2 for positron spinors. The summation over the spins yield

$$\sum_{\text{spins}} u(\boldsymbol{p})\bar{u}(\boldsymbol{p}) = \sum_{1,2} u_i(\boldsymbol{p})\bar{u}_i(\boldsymbol{p}) = \not{p} + m; \tag{2.69}$$

$$\sum_{\text{spins}} v(\boldsymbol{p})\bar{v}(\boldsymbol{p}) = \sum_{1,2} v_i(\boldsymbol{p})\bar{v}_i(\boldsymbol{p}) = \not{p} - m. \tag{2.70}$$

2.5.3 Sum over spin states

Let us now reconsider the problem of summing over spin states using Feynman's notation. The square of the matrix element (2.31) can be evaluated using Feynman's projection operator for positive energy states.

$$\begin{aligned} |T_{fi}|^2 &= \frac{1}{2}\sum_{if}(\bar{\psi}_f\mathcal{O}\psi_i)(\bar{\psi}_f\mathcal{O}\psi_i)^\dagger \\ &= \frac{1}{2}\sum_{if}(\bar{\psi}_f\mathcal{O}\psi_i)(\psi_i^\dagger\mathcal{O}^\dagger\bar{\psi}_f^\dagger) \\ &= \frac{1}{2}\sum_{if}(\bar{\psi}_f\mathcal{O}\psi_i)(\bar{\psi}_i\gamma_0\mathcal{O}^\dagger\gamma_0\psi_f) \\ &= \frac{1}{2}\sum_{if}(\bar{\psi}_f\mathcal{O}\psi_i)(\bar{\psi}_i\tilde{\mathcal{O}}\psi_f), \end{aligned} \tag{2.71}$$

where $\sum_{if}$ denotes the summation over the positive energy spin states of the incident and scattered particle and $\tilde{\mathcal{O}}$ stands for

$$\tilde{\mathcal{O}} = \gamma_0\mathcal{O}^\dagger\gamma_0. \tag{2.72}$$

Replacing $\sum_i \psi_i\bar{\psi}_i$ by $(\not{p}_i+m)$ which is the projection operator for positive energy states, we obtain

$$|T_{fi}|^2 = \frac{1}{2}\sum_{f}(\bar{\psi}_f\mathcal{O}(\not{p}_i + m)\tilde{\mathcal{O}}\psi_f). \tag{2.73}$$

Using the algebra of matrix multiplication, we get

$$\begin{aligned} |T_{fi}|^2 &= \frac{1}{2}\sum_{f}\sum_{\rho,\lambda,\rho',\lambda'}(\bar{\psi}_f)_\rho(\mathcal{O})_{\rho\lambda}(\not{p}_i + m)_{\lambda\rho'}(\tilde{\mathcal{O}})_{\rho'\lambda'}(\psi_f)_{\lambda'} \\ &= \frac{1}{2}\sum_{f}\sum_{\rho,\lambda,\rho',\lambda'}(\mathcal{O})_{\rho\lambda}(\not{p}_i + m)_{\lambda\rho'}(\tilde{\mathcal{O}})_{\rho'\lambda'}(\psi_f\bar{\psi}_f)_{\lambda'\rho}. \end{aligned} \tag{2.74}$$

Summing over f over the two positive energy spin states and replacing $\sum_f \psi_f \bar{\psi}_f$ by $(\not{p}_f + m)$, which is the projection operator for the positive energy states, in Eq. (2.74); we get

$$|T_{fi}|^2 = \frac{1}{2}\mathrm{Tr}\left(\mathcal{O}\,(\not{p}_i + m)\,\tilde{\mathcal{O}}\,(\not{p}_f + m)\right). \tag{2.75}$$

The transition probability per unit time is given by Fermi's golden rule

$$\text{Transition rate} = 2\pi(\Pi N)^{-1}|T_{fi}|^2\,\rho_f, \tag{2.76}$$

where ΠN denotes the normalization factor $2E$ for each of the initial and final particles and ρ_f is the density of states for the final particle. The cross section is the transition rate per unit incident flux.

2.6 A Consistency Check

We have deduced two different expressions (2.32) and (2.75), for the square of the transition amplitude $|T_{fi}|^2$, one using the Dirac matrices and the other using Feynman's notation. They must be equivalent. To show this, let us start with the projection operator Λ^+ for positive energy states as given by Eq. (2.23). (Please note that in natural units $W = E_p$.)

$$\Lambda^+ \;=\; \frac{1}{2}\left(I + \frac{\boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m}{E_p}\right) = \frac{1}{2E_p}(E_p + \boldsymbol{\alpha}\cdot\boldsymbol{p} + \beta m). \tag{2.77}$$

Multiply Eq. (2.77) by $\beta^2 = I$ from the right to obtain

$$\begin{aligned}
\Lambda^+ &= \frac{1}{2E_p}(\beta E_p + \boldsymbol{\alpha}\cdot\boldsymbol{p}\,\beta + m)\beta \\
&= \frac{1}{2E_p}(\beta E_p - \beta\boldsymbol{\alpha}\cdot\boldsymbol{p} + m)\beta \\
&= \frac{1}{2E_p}(\gamma_0 E_p - \boldsymbol{\gamma}\cdot\boldsymbol{p} + m)\beta = \frac{1}{2E_p}(\not{p} + m)\beta.
\end{aligned} \tag{2.78}$$

Substituting the expression (2.78) for Λ^+, into Eq. (2.32) and remembering that $\beta \equiv \gamma_0$ and $\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA)$, we get

$$\begin{aligned}
|T_{fi}|^2 &= \frac{1}{8E_iE_f}\mathrm{Tr}\left(\gamma_0\mathcal{O}(\not{p}_i + m)\gamma_0\mathcal{O}^\dagger\gamma_0(\not{p}_f + m)\gamma_0\right) \\
&= \frac{1}{8E_iE_f}\mathrm{Tr}\left(\mathcal{O}(\not{p}_i + m)\tilde{\mathcal{O}}(\not{p}_f + m)\right).
\end{aligned} \tag{2.79}$$

This is identical with Eq. (2.75) except for the additional factor $1/(4E_iE_f)$ which we include, in Feynman's formalism, as normalization factor $(\Pi N)^{-1}$, as indicated in Eq. (2.76).

2.7 Algebra of γ Matrices

The square of the transition matrix element given by Eq. (2.75) involves the trace of a product of γ matrices. So, it will be fruitful to study the γ matrices and their properties for evaluating he traces. The γ matrices obey the following relations:

$$\begin{aligned} &\gamma_0^2 = 1, \qquad \gamma_x^2 = \gamma_y^2 = \gamma_z^2 = -1, \\ &\gamma_0\gamma_{x,y,z} + \gamma_{x,y,z}\gamma_0 = 0, \\ &\gamma_x\gamma_y + \gamma_y\gamma_x = 0, \qquad \gamma_y\gamma_z + \gamma_z\gamma_y = 0, \qquad \gamma_z\gamma_x + \gamma_x\gamma_z = 0. \end{aligned} \tag{2.80}$$

Using a unified notation, Eq. (2.80) can be written as

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}, \tag{2.81}$$

where $g_{\mu\nu}$ is a metric defined by

$$g_{\mu\nu} = \begin{cases} 0, & \mu \neq \nu \\ +1, & \mu = \nu = 0 \\ -1, & \mu = \nu = x, y, z. \end{cases} \tag{2.82}$$

Besides, the matrix γ_0 is Hermitian whereas the matrices $\gamma_x, \gamma_y, \gamma_z$ are anti-Hermitian.

$$\gamma_0^\dagger = \gamma_0; \qquad \gamma_k^\dagger = -\gamma_k, \ k = x, y, z. \tag{2.83}$$

It is convenient to define[5] a matrix γ_5 which occurs frequently.

$$\gamma_5 = \gamma_0\gamma_x\gamma_y\gamma_z = -i \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}. \tag{2.84}$$

It is easy to verify that

$$\gamma_5^\dagger = -\gamma_5; \quad \gamma_5^2 = -1; \quad \gamma_5\gamma_\mu + \gamma_\mu\gamma_5 = 0. \tag{2.85}$$

Following Feynman, we can define $\not{a}$ as follows:

$$\not{a} = a_0\gamma_0 - a_x\gamma_x - a_y\gamma_y - a_z\gamma_z. \tag{2.86}$$

[5]Feynman's definition of $\gamma_5 = \gamma_x\gamma_y\gamma_z\gamma_0$ differs by a change in sign.

It can be shown that

$$\begin{aligned}
\not{d}\gamma_5 &= -\gamma_5\not{d}; && (2.87)\\
\not{d}\not{b} &= -\not{b}\not{d} + 2\,\mathbf{a}\cdot\mathbf{b} \quad (\mathbf{a}\cdot\mathbf{b} = a_\mu b_\mu); && (2.88)\\
\gamma_x\not{d}\gamma_x &= \not{d} + 2a_x\gamma_x; && (2.89)\\
\gamma_\mu\not{d}\gamma_\mu &= -2\not{d}; && (2.90)\\
\gamma_\mu\not{d}\not{b}\gamma_\mu &= 4\,\mathbf{a}\cdot\mathbf{b}; && (2.91)\\
\gamma_\mu\not{d}\not{b}\not{c}\gamma_\mu &= -2\not{c}\not{b}\not{d}. && (2.92)
\end{aligned}$$

It is important to recall the elementary properties of traces,

$$\begin{aligned}
&\mathrm{Tr}\,(ABC) = \mathrm{Tr}\,(BCA) = \mathrm{Tr}\,(CAB),\\
&\mathrm{Tr}\,(A+B) = \mathrm{Tr}\,A + \mathrm{Tr}\,B,
\end{aligned} \qquad (2.93)$$

for evaluating the traces involving a product of γ matrices. It is known that the trace of a γ matrix is zero.

$$\mathrm{Tr}\,\gamma_\mu = 0 \qquad (\mu = 0, x, y, z). \qquad (2.94)$$

Also the trace of an odd number of γ matrices vanishes. To prove this, we start with the relation (2.85) which is equivalent to

$$\gamma_5\gamma_\mu(\gamma_5)^{-1} = -\gamma_\mu. \qquad (2.95)$$

It follows that

$$\gamma_5\gamma_{\mu_1}\gamma_{\mu_2}\cdots\gamma_{\mu_n}(\gamma_5)^{-1} = (-1)^n\gamma_{\mu_1}\gamma_{\mu_2}\cdots\gamma_{\mu_n}. \qquad (2.96)$$

Taking the trace of both sides of Eq. (2.96) and using the elementary property of the trace that $\mathrm{Tr}\,(ABC) = \mathrm{Tr}\,(BCA)$, we obtain immediately that

$$\mathrm{Tr}(\gamma_{\mu_1}\cdots\gamma_{\mu_n}) = (-1)^n\mathrm{Tr}(\gamma_{\mu_1}\cdots\gamma_{\mu_n}). \qquad (2.97)$$

Equation (2.97) implies that the the trace of an odd number of gamma matrices vanishes.

If n is even, it is always possible to reduce it to $n-2$ factors. For example,

$$\begin{aligned}
\mathrm{Tr}(\gamma_\mu\gamma_\nu) &= \mathrm{Tr}\,(\gamma_\nu\gamma_\mu)\\
&= \tfrac{1}{2}\,\mathrm{Tr}\,(\gamma_\nu\gamma_\mu + \gamma_\mu\gamma_\nu), && \text{since Tr(AB)=Tr(BA)}\\
&= g_{\mu\nu}\,\mathrm{Tr}\,I, && \text{using Eq. (2.81)}\\
&= 4\,g_{\mu\nu}. && (2.98)
\end{aligned}$$

In a similar way, it can be shown that

$$\mathrm{Tr}\,(\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\lambda) = 4g_{\mu\lambda}g_{\nu\rho} - 4g_{\lambda\nu}g_{\rho\mu} + 4g_{\lambda\rho}g_{\mu\nu}. \tag{2.99}$$

The following traces which occur frequently are given[6].

$$\begin{aligned} \mathrm{Tr}\,(\not{a}\not{b}) &= \tfrac{1}{2}\mathrm{Tr}\,(\not{a}\not{b} + \not{b}\not{a}) = \mathrm{Tr}\,(\mathbf{a}\cdot\mathbf{b}) = 4\,\mathbf{a}\cdot\mathbf{b}, \\ \mathrm{Tr}\,(\not{a}\not{b}\not{c}) &= 0. \end{aligned} \tag{2.100}$$

Also, it can be shown that the trace of a product of γ matrices is the same as the product of γ matrices taken in the reverse order.

$$\mathrm{Tr}\,(\not{a}\not{b}\not{c}\not{d}) = \mathrm{Tr}\,(\not{d}\not{c}\not{b}\not{a}). \tag{2.101}$$

2.8 An Illustrative Example

We shall illustrate the foregoing discussion by evaluating the trace in Eq. (2.75) when the transition operator is given by

$$\mathcal{O} = \gamma_0. \tag{2.102}$$

Substituting the operator $\mathcal{O}$ in Eq. (2.75), we obtain

$$|T_{fi}|^2 = \frac{1}{2}\mathrm{Tr}\left(\gamma_0(\not{p}_i + m)\,\tilde{\gamma}_0\,(\not{p}_f + m)\right), \tag{2.103}$$

where $\not{p} = \gamma_0 E - \boldsymbol{\gamma}\cdot\boldsymbol{p}$. Since $\tilde{\gamma}_0 = \gamma_0$ and the trace of a product of an odd number of γ matrices vanishes,

$$\begin{aligned} \mathrm{Tr}\,\left(\gamma_0(\not{p}_i + m)\,\gamma_0\,(\not{p}_f + m)\right) &= \mathrm{Tr}\,\left(\gamma_0\not{p}_i\,\gamma_0\not{p}_f + m^2\right) \\ &= \mathrm{Tr}\,\left((-\not{p}_i\gamma_0 + 2E_i)\gamma_0\not{p}_f + m^2\right) \\ &= \mathrm{Tr}\,\left(-\not{p}_i\not{p}_f + 2E_i\gamma_0\not{p}_f + m^2\right) \\ &= -4\,p_i\cdot p_f + 8E_iE_f + 4m^2. \end{aligned} \tag{2.104}$$

Equations (2.98) and (2.100) have been used in deducing the last step in the above equation. Expanding the scalar product of the four-vectors $\mathbf{p}_i\cdot\mathbf{p}_f = E_iE_f - \boldsymbol{p}_i\cdot\boldsymbol{p}_f$ and rearranging, we get

$$\mathrm{Tr}\left(\gamma_0(\not{p}_i + m)\,\gamma_0\,(\not{p}_f + m)\right) = 4E_iE_f + 4p_ip_f\cos\theta + 4m^2, \tag{2.105}$$

where θ denotes the angle between the two vectors $\boldsymbol{p}_i$ and $\boldsymbol{p}_f$.

[6] We use bold upright letters $\mathbf{a}, \mathbf{b}$ to denote the four-vectors to distinguish them from the three-vectors for which we use bold italics $\boldsymbol{a}, \boldsymbol{b}$.

Review Questions

2.1 Write down the Dirac equation for a free particle using the α and β matrices and obtain its solutions. How many solutions are there and how are they interpreted? Discuss their orthogonal and closure properties.

2.2 Construct the projection operators for the positive and negative energy states of the Dirac Hamiltonian and demonstrate how they pick up only positive or negative energy states.

2.3 Obtain the Dirac equation in Feynman's notation using γ matrices and find its free particle solutions. Show that the positive eigenvalue solutions coincide with those of Dirac whereas the negative eigenvalue solutions differ from the negative energy solutions of Dirac. How will you account for this discrepancy?

2.4 A Dirac particle is scattered from an initial state to a final state due to an interaction represented by the operator $\mathcal{O}$. If the initial and final state spins are not observed, how will you calculate the transition probability by summing over the final spin states and averaging over the initial spin states.

2.5 Write down the gamma matrices and discuss their properties. Show that the trace of a product of an odd number of γ matrices vanishes. Evaluate (a) $\mathrm{Tr}\,(\gamma_\mu\gamma_\nu)$ and (b) $\mathrm{Tr}\,(\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\lambda)$.

Problems

2.1 Show that the positive eigenvalue states ψ_p and the negative eigenvalue states ψ_n of the Dirac equation, given by

$$\psi_p = \sqrt{E_p + m}\begin{bmatrix} \chi_\pm \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E+m}\chi_\pm \end{bmatrix}, \qquad \psi_n = \sqrt{E_p + m}\begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_\pm \\ \chi_\pm \end{bmatrix}.$$

are normalized to yield

$$\psi_p^\dagger\psi_p = 2E_p, \qquad \psi_n^\dagger\psi_n = 2E_p,$$

and

$$\bar{\psi}_p\psi_p = 2m, \qquad \bar{\psi}_n\psi_n = -2m.$$

2.2 Show that, in Feynman's notation, a sum over spins of positive eigenvalue states and a sum over spins of negative eigenvalue states, done separately, yield

$$\sum_{\text{spins}} \psi_p\bar{\psi}_p = \not{p} + m; \qquad \sum_{\text{spins}} \psi_n\bar{\psi}_n = \not{p} - m.$$

Show that they can be considered as unnormalized projection operators for positive and negative eigenvalue states of Dirac equation written in Feynman's notation. Hence obtain the normalized projection operators.

2.3 Given the transition operator $\mathcal{O} = \boldsymbol{\gamma} \cdot \boldsymbol{A}$, where $\boldsymbol{A}$ is a vector in the three-dimensional space and not an operator, calculate the transition probability for an electron if the initial and final spin states are not observed.

2.4 Given the transition operator $\mathcal{O} = \gamma_\mu J_\mu$ where J_μ is a four-vector current, calculate the transition probability for an electron if the initial and final spin states are not observed.

2.5 Show that if the operator $\mathcal{O} = \not{a}\not{b}\not{c}$, then $\tilde{\mathcal{O}} = \gamma_0 \mathcal{O}^\dagger \gamma_0 = \not{c}\not{b}\not{a}$.

Solutions to Problems

2.1 First let us show that the Pauli spin vectors are orthonormal.

$$\chi_+^\dagger \chi_+ = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1. \qquad \chi_+^\dagger \chi_- = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0.$$
$$\chi_-^\dagger \chi_- = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 1.$$

Altogether, there are four solutions for the Dirac equation. Let us denote the two solutions that correspond to positive eigenvalue, one with spin up and the other with spin down by $\psi_{p\uparrow}, \psi_{p\downarrow}$. The other two solutions that correspond to negative eigenvalue relating to the two spin orientations, spin-up and spin-down are denoted by $\psi_{n\uparrow}, \psi_{n\downarrow}$. Using the orthonormal property of the Pauli spin vectors, it is shown below by direct matrix multiplication that

$$\psi_{p\uparrow}^\dagger \psi_{p\uparrow} = 2E_p. \qquad \psi_{n\uparrow}^\dagger \psi_{n\uparrow} = 2E_p.$$

$$\begin{aligned}
\psi_{p\uparrow}^\dagger \psi_{p\uparrow} &= (E_p + m) \begin{bmatrix} \chi_+^\dagger & \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+^\dagger \end{bmatrix} \begin{bmatrix} \chi_+ \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+ \end{bmatrix} \\
&= (E_p + m) \left\{ 1 + \frac{(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2}{(E_p+m)^2} \right\} \\
&= (E_p + m) \left\{ 1 + \frac{p^2}{(E_p+m)^2} \right\} \\
&= (E_p + m) \left\{ 1 + \frac{E_p^2 - m^2}{(E_p+m)^2} \right\}, \\
&\qquad \text{using the relation } E_p^2 = p^2 + m^2 \\
&= 2E_p.
\end{aligned}$$

$$\begin{aligned}
\psi^\dagger_{n\uparrow}\psi_{n\uparrow} &= (E_p+m)\begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+^\dagger & \chi_+^\dagger \end{bmatrix}\begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+ \\ \chi_+ \end{bmatrix} \\
&= (E_p+m)\left\{\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2}{(E_p+m)^2}+1\right\} \\
&= (E_p+m)\left\{\frac{p^2}{(E_p+m)^2}+1\right\} \\
&= 2E_p.
\end{aligned}$$

In a similar way, it can be shown that $\psi^\dagger_{p\downarrow}\psi_{p\downarrow} = 2E_p$. $\psi^\dagger_{n\downarrow}\psi_{n\downarrow} = 2E_p$. Further,

$$\begin{aligned}
\bar\psi_{p\uparrow}\psi_{p\uparrow} &= \psi^\dagger_{p\uparrow}\gamma_0\psi_{p\uparrow} \\
&= (E_p+m)\begin{bmatrix} \chi_+^\dagger & \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+^\dagger \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} \chi_+ \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+ \end{bmatrix} \\
&= (E_p+m)\begin{bmatrix} \chi_+^\dagger & \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+^\dagger \end{bmatrix}\begin{bmatrix} \chi_+ \\ -\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+ \end{bmatrix} \\
&= (E_p+m)\left\{1-\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2}{(E_p+m)^2}\right\} \\
&= (E_p+m)\left\{1-\frac{p^2}{(E_p+m)^2}\right\} \\
&= (E_p+m)\left\{1-\frac{E_p^2-m^2}{(E_p+m)^2}\right\}, \\
&\qquad \text{using the relation } E_p^2 = p^2+m^2 \\
&= 2m.
\end{aligned}$$

$$\begin{aligned}
\bar\psi_{n\uparrow}\psi_{n\uparrow} &= \psi^\dagger_{n\uparrow}\gamma_0\psi_{n\uparrow} \\
&= (E_p+m)\begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+^\dagger & \chi_+^\dagger \end{bmatrix}\begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}\begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+ \\ \chi_+ \end{bmatrix} \\
&= (E_p+m)\begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+^\dagger & \chi_+^\dagger \end{bmatrix}\begin{bmatrix} \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+ \\ -\chi_+ \end{bmatrix} \\
&= (E_p+m)\left\{\frac{(\boldsymbol{\sigma}\cdot\boldsymbol{p})^2}{(E_p+m)^2}-1\right\} \\
&= (E_p+m)\left\{\frac{p^2}{(E_p+m)^2}-1\right\} = -2m.
\end{aligned}$$

Similarly, it follows that $\bar\psi_{p\downarrow}\psi_{p\downarrow} = 2m$; $\bar\psi_{n\downarrow}\psi_{n\downarrow} = -2m$.

2.2 The quantities $\psi_p\bar{\psi}_p$ and $\psi_n\bar{\psi}_n$ are matrices.
It can be easily checked that $\chi_+\chi_+^\dagger + \chi_-\chi_-^\dagger$ is a unit matrix.

$$\chi_+\chi_+^\dagger + \chi_-\chi_-^\dagger = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$\begin{aligned}
\sum_{\text{spins}} \psi_p\bar{\psi}_p &= (E_p+m)\left\{ \begin{bmatrix} \chi_+ \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+ \end{bmatrix} \begin{bmatrix} \chi_+^\dagger & \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_+^\dagger \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \right. \\
&\qquad \left. + \begin{bmatrix} \chi_- \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_- \end{bmatrix} \begin{bmatrix} \chi_-^\dagger & \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m}\chi_-^\dagger \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \right\} \\
&= (E_p+m) \begin{bmatrix} 1 & \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m} \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m} & \frac{p^2}{(E_p+m)^2} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \\
&= (E_p+m) \begin{bmatrix} 1 & -\frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m} \\ \frac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p+m} & -\frac{p^2}{(E_p+m)^2} \end{bmatrix} \\
&= \begin{bmatrix} E_p+m & -\boldsymbol{\sigma}\cdot\boldsymbol{p} \\ \boldsymbol{\sigma}\cdot\boldsymbol{p} & -(E_p-m) \end{bmatrix} \\
&= \not{p}+m,
\end{aligned}$$

since

$$\begin{aligned}
\not{p}+m &= \gamma_0 E_p - \boldsymbol{\gamma}\cdot\boldsymbol{p} + mI \\
&= \begin{bmatrix} E_p & 0 \\ 0 & E_p \end{bmatrix} - \begin{bmatrix} 0 & \boldsymbol{\sigma}\cdot\boldsymbol{p} \\ -\boldsymbol{\sigma}\cdot\boldsymbol{p} & 0 \end{bmatrix} + m\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}$$

Similarly, it can be shown that

$$\sum_{\text{spins}} \psi_n\bar{\psi}_n = \not{p}-m.$$

The operators $\not{p}+m$ and $\not{p}-m$ are called projection operators since they pick out selectively the positive energy states and negative energy states of the electron.

$$\begin{aligned}
&(\not{p}+m)\psi_p = 2m\psi_p; \qquad && (\not{p}+m)\psi_n = 0. \\
&(\not{p}-m)\psi_p = 0; && (\not{p}-m)\psi_n = -2m\psi_n.
\end{aligned}$$

These projection operators are not normalized. The normalized projection operators are $\frac{\not{p}+m}{2m}$ for positive energy states and $\frac{m-\not{p}}{2m}$ for negative energy states.

2.3 From Eq. (2.75), we have

$$|T_{fi}|^2 = \frac{1}{2}\mathrm{Tr}\left\{\mathcal{O}\left(\not{p}_i + m\right)\tilde{\mathcal{O}}\left(\not{p}_f + m\right)\right\},$$

where

$$\tilde{\mathcal{O}} = \gamma_0 \mathcal{O}^\dagger \gamma_0.$$

The transition operator $\mathcal{O}$, in the present case is $\boldsymbol{\gamma}\cdot\boldsymbol{A}$. Substituting it, we shall write down the product of operators $\{\cdots\}$.

$$\begin{aligned}\{\cdots\} &= \gamma_k A_k \left(\gamma_\mu (p_i)_\mu + m\right)\gamma_0\gamma_l^\dagger A_l^*\gamma_0\left(\gamma_\nu (p_f)_\nu + m\right)\\ &= \gamma_k A_k \left(\gamma_\mu (p_i)_\mu + m\right)\gamma_l A_l^*\left(\gamma_\nu (p_f)_\nu + m\right).\end{aligned}$$

We have used above the relations $\gamma_l^\dagger = -\gamma_l$, $\gamma_0\gamma_k = -\gamma_k\gamma_0$ and $\gamma_0^2 = 1$. Since the trace of a product of odd number of γ matrices is zero,

$$\mathrm{Tr}\left\{\cdots\right\} = \mathrm{Tr}\{\gamma_k\gamma_\mu\gamma_l\gamma_\nu A_k A_l^*(p_i)_\mu (p_f)_\nu + \gamma_k\gamma_l A_k A_l^* m^2\}.$$

The indices k and l denote the components of a three vector and the indices μ and ν denote the components of a four-vector. We have earlier evaluated the traces of even number of γ matrices.

$$\begin{aligned}&\mathrm{Tr}(\gamma_k\gamma_\mu\gamma_l\gamma_\nu) = 4g_{k\mu}g_{l\nu} - 4g_{kl}g_{\mu\nu} + 4g_{k\nu}g_{\mu l};\\ &\mathrm{Tr}(\gamma_k\gamma_l) = 4g_{kl}.\end{aligned}$$

Using the above results,

$$\begin{aligned}|T_{fi}|^2 &= \frac{1}{2}\{4g_{k\mu}g_{l\nu}(p_i)_\mu A_k (p_f)_\nu A_l^* - 4g_{kl}g_{\mu\nu}A_k A_l^*(p_i)_\mu (p_f)_\nu\\ &\quad + 4g_{k\nu}g_{\mu l}(p_f)_\nu A_k (p_i)_\mu A_l^* + 4g_{kl}A_k A_l^* m^2\}\\ &= \frac{1}{2}\left\{4(\boldsymbol{p}_i\cdot\boldsymbol{A})(\boldsymbol{p}_f\cdot\boldsymbol{A}^*) + 4(\boldsymbol{A}\cdot\boldsymbol{A}^*)(\mathbf{p}_i\cdot\mathbf{p}_f)\right.\\ &\quad \left. + 4(\boldsymbol{p}_f\cdot\boldsymbol{A})(\boldsymbol{p}_i\cdot\boldsymbol{A}^*) - 4m^2\boldsymbol{A}\cdot\boldsymbol{A}^*\right\}\\ &= 2\left\{(\boldsymbol{p}_i\cdot\boldsymbol{A})(\boldsymbol{p}_f\cdot\boldsymbol{A}^*) + (\boldsymbol{p}_f\cdot\boldsymbol{A})(\boldsymbol{p}_i\cdot\boldsymbol{A}^*)\right.\\ &\quad \left. + (\boldsymbol{A}\cdot\boldsymbol{A}^*)(\mathbf{p}_i\cdot\mathbf{p}_f - m^2)\right\},\end{aligned}$$

where $\mathbf{p}_i\cdot\mathbf{p}_f$ denotes the scalar product of four vectors, defined by

$$\mathbf{p}_i\cdot\mathbf{p}_f = E_iE_f - \boldsymbol{p}_i\cdot\boldsymbol{p}_f.$$

2.4 From Eq. (2.75), we have

$$|T_{fi}|^2 = \frac{1}{2}\mathrm{Tr}\left\{\mathcal{O}\left(\not{p}_i + m\right)\tilde{\mathcal{O}}\left(\not{p}_f + m\right)\right\},$$

where

$$\tilde{\mathcal{O}} = \gamma_0 \mathcal{O}^\dagger \gamma_0.$$

The transition operator $\mathcal{O}$, in the present case is $\gamma_\mu J_\mu$ where J_μ is a four-vector current.

$$\tilde{\mathcal{O}} = \gamma_0 \gamma_\mu^\dagger \gamma_0 J_\mu^* = \gamma_\mu J_\mu^*.$$

Substituting them, we shall write down the product of operators $\{\cdots\}$.

$$\{\cdots\} \quad = \quad \left[\gamma_\mu (\not{p}_i + m)\gamma_\nu (\not{p}_f + m)\right] J_\mu J_\nu^*$$

Since the trace of a product of odd number of γ matrices is zero, we find

$$\begin{aligned}
|T_{fi}|^2 &= \frac{1}{2}\left[\mathrm{Tr}(\gamma_\mu \not{p}_i \gamma_\nu \not{p}_f) J_\mu J_\nu^* + m^2 \gamma_\mu \gamma_\nu J_\mu J_\nu^*\right] \\
&= \frac{1}{2}\left[\mathrm{Tr}(\gamma_\mu \gamma_\rho \gamma_\nu \gamma_\lambda)(p_i)_\rho (p_f)_\lambda J_\mu J_\nu^* + m^2 \mathrm{Tr}(\gamma_\mu \gamma_\nu) J_\mu J_\nu^*\right] \\
&= \frac{1}{2}\left[\{4g_{\mu\lambda}g_{\nu\rho} - 4g_{\mu\nu}g_{\lambda\rho} + 4g_{\mu\rho}g_{\nu\lambda}\}(p_i)_\rho (p_f)_\lambda J_\mu J_\nu^*\right. \\
&\quad \left. + 4m^2 (g_{\mu\nu}) J_\mu J_\nu^*\right] \\
&= \frac{1}{2}\left[4(\mathbf{p}_f \cdot \mathbf{J})(\mathbf{p}_i \cdot \mathbf{J}^*) - 4(\mathbf{p}_i \cdot \mathbf{p}_f)(\mathbf{J} \cdot \mathbf{J}^*)\right. \\
&\quad \left. + 4(\mathbf{p}_i \cdot \mathbf{J})(\mathbf{p}_f \cdot \mathbf{J}^*) + 4m^2 (\mathbf{J} \cdot \mathbf{J}^*)\right] \\
&= 2\left[(\mathbf{p}_i \cdot \mathbf{J})(\mathbf{p}_f \cdot \mathbf{J}^*) + (\mathbf{p}_f \cdot \mathbf{J})(\mathbf{p}_i \cdot \mathbf{J}^*) + (m^2 - \mathbf{p}_i \cdot \mathbf{p}_f)\mathbf{J} \cdot \mathbf{J}^*\right],
\end{aligned}$$

where $(\mathbf{p}_i \cdot \mathbf{J}), \ldots$ are scalar products of four-vectors.

2.5 Given

$$\mathcal{O} = \not{a}\not{b}\not{c},$$

then

$$\begin{aligned}
\tilde{\mathcal{O}} &= \gamma_0 \mathcal{O}^\dagger \gamma_0 \\
&= \gamma_0 (\not{a}\not{b}\not{c})^\dagger \gamma_0 \\
&= \gamma_0 (\gamma_\mu a_\mu \gamma_\nu b_\nu \gamma_\lambda c_\lambda)^\dagger \gamma_0 \\
&= \gamma_0 \left\{(\gamma_\mu \gamma_\nu \gamma_\lambda)^\dagger a_\mu b_\nu c_\lambda\right\} \gamma_0 \\
&= \gamma_0 (\gamma_\lambda^\dagger \gamma_\nu^\dagger \gamma_\mu^\dagger a_\mu b_\nu c_\lambda) \gamma_0 \\
&= \gamma_0 \left\{(\gamma_0 c_0 + \boldsymbol{\gamma} \cdot \boldsymbol{c})(\gamma_0 b_0 + \boldsymbol{\gamma} \cdot \boldsymbol{b})(\gamma_0 a_0 + \boldsymbol{\gamma} \cdot \boldsymbol{a})\right\} \gamma_0,
\end{aligned}$$

since $\gamma_0^\dagger = \gamma_0$ and $\gamma_k^\dagger = -\gamma_k,\ k = 1, 2, 3$. Switching the γ_0 in the front to the rear side using the commutation relations $\gamma_0 \boldsymbol{\gamma} = -\boldsymbol{\gamma}\gamma_0$ and noting that $\gamma_0^2 = 1$, we get

$$\begin{aligned}
\tilde{\mathcal{O}} &= (\gamma_0 c_0 - \boldsymbol{\gamma} \cdot \boldsymbol{c})(\gamma_0 b_0 - \boldsymbol{\gamma} \cdot \boldsymbol{b})(\gamma_0 a_0 - \boldsymbol{\gamma} \cdot \boldsymbol{a}) \\
&= \not{c}\not{b}\not{a}
\end{aligned}$$

Chapter 3

The Propagation Kernel and Feynman Diagrams

In this chapter, we shall introduce the concept of propagation kernel and discuss how Feynman modified it in such a way that the positive energy states in relativistic Dirac theory are propagated forward in time whereas the negative energy states are propagated backwards in time. This new view-point has completely revolutionized the study of quantum electrodynamics and enabled the visualization of quantum electrodynamic processes by means of Feynman diagrams. It is said that more than Feynman, it was Freeman Dyson who has popularized the Feynman Diagrams and showed that Feynman's approach is equivalent to the field theoretical method.

For the study of a given electrodynamic process, one has to draw all possible Feynman diagrams for the process and calculate the transition probability by formulating a set of rules known as Feynman rules.

3.1 The Propagation Kernel

3.1.1 In non-relativistic Schrödinger theory

First let us consider the non-relativistic Schrödinger equation

$$i\hbar\frac{\partial\Psi(\boldsymbol{x},t)}{\partial t} = H\Psi(\boldsymbol{x},t), \tag{3.1}$$

and introduce the concept of propagation kernel. It is an extension of the Green's function technique[1], introduced in the theory of scattering. Equation (3.1) gives the infinitesimal change in the wave function Ψ due to a small increment in time. To obtain $\Psi(\boldsymbol{x}_2, t_2)$ from $\Psi(\boldsymbol{x}_1, t_1)$, one has to go through a series of successive incremental steps but Feynman chose to take a great leap forward with the help of a propagation kernel. The propagation kernel $K(\boldsymbol{x}_2 t_2; \boldsymbol{x}_1 t_1)$ describes the evolution of the state function $\Psi(\boldsymbol{x}, t)$ and it is defined by

$$\Psi(\boldsymbol{x}_2, t_2) = \int d^3x_1 \, K(\boldsymbol{x}_2 t_2 : \boldsymbol{x}_1 t_1)\Psi(\boldsymbol{x}_1, t_1), \qquad t_2 \geq t_1. \tag{3.2}$$

First let us assume that H is a time independent operator with eigenfunctions $\psi_n(\boldsymbol{x})$ and eigenvalues E_n, such that

$$H\psi_n(\boldsymbol{x}) = E_n\psi_n(\boldsymbol{x}). \tag{3.3}$$

The eigenfunctions are orthonormal and obey the closure property.

$$\int d^3\boldsymbol{x}\psi_n^*(\boldsymbol{x})\psi_m(\boldsymbol{x}) = \delta_{mn}; \tag{3.4}$$

$$\sum_n \psi_n(\boldsymbol{x})\psi_n(\boldsymbol{x}') = \delta(\boldsymbol{x} - \boldsymbol{x}'). \tag{3.5}$$

We can now expand $\Psi(\boldsymbol{x}_1, t_1)$ in terms of the complete set of orthonormal eigenfunctions $\psi_n(\boldsymbol{x})$.

$$\Psi(\boldsymbol{x}_1, t_1) = \sum_n c_n\psi_n(\boldsymbol{x}_1)e^{-iE_n t_1/\hbar}. \tag{3.6}$$

From Eq. (3.6), the expansion coefficients c_n can be obtained by using the orthonormal and closure properties of $\psi_n(\boldsymbol{x})$.

$$c_n = \int \psi_n^*(\boldsymbol{x}_1)\Psi(\boldsymbol{x}_1, t_1)e^{iE_n t_1/\hbar}d^3x_1. \tag{3.7}$$

Using these expansion coefficients, the state vector $\Psi(\boldsymbol{x}_2, t_2)$ can be written as

$$\Psi(\boldsymbol{x}_2, t_2) = \sum_n c_n\psi_n(\boldsymbol{x}_2)e^{-iE_n t_2/\hbar} \tag{3.8}$$

$$= \sum_n \int \psi_n^*(\boldsymbol{x}_1)\Psi(\boldsymbol{x}_1, t_1)\psi_n(\boldsymbol{x}_2)e^{-iE_n(t_2-t_1)/\hbar}\, d^3\boldsymbol{x}_1. \tag{3.9}$$

[1]V. Devanathan, *Quantum Mechanics*, (Ch.9, Theory of Scattering) Narosa Publishers, New Delhi (2005)

Comparing Eqs. (3.2) and (3.9), we obtain an expression for the propagation kernel $K(\boldsymbol{x}_2 t_2; \boldsymbol{x}_1 t_1)$.

$$\begin{aligned} K(\boldsymbol{x}_2 t_2; \boldsymbol{x}_1 t_1) &= \sum_n \psi_n(\boldsymbol{x}_2)\psi_n^*(\boldsymbol{x}_1) e^{-iE_n(t_2-t_1)/\hbar}, \qquad t_2 > t_1 \\ &= \sum_n \chi_n(\boldsymbol{x}_2, t_2)\chi_n^*(\boldsymbol{x}_1, t_1)\theta(t_2 - t_1), \end{aligned} \tag{3.10}$$

where

$$\chi_n(\boldsymbol{x}, t) = \psi_n(\boldsymbol{x}) e^{-iE_n t/\hbar}, \tag{3.11}$$

$$\theta(t_2 - t_1) = \begin{cases} 1, & \text{if } t_2 > t_1 \\ 0, & \text{if } t_2 < t_1 \end{cases}. \tag{3.12}$$

It is now possible to obtain a differential equation which the propagation kernel obeys. Since $i\hbar \frac{\partial}{\partial t_2}\chi_n(\boldsymbol{x}_2, t_2) = H\chi_n(\boldsymbol{x}_2, t_2)$, we obtain

$$\begin{aligned} &\left\{ i\hbar \frac{\partial}{\partial t_2} - H(\boldsymbol{x}_2) \right\} K(\boldsymbol{x}_2 t_2; \boldsymbol{x}_1 t_1) \\ &= i\hbar \sum_n \chi_n(\boldsymbol{x}_2, t_2)\chi_n^*(\boldsymbol{x}_1, t_1) \frac{\partial}{\partial t_2}\theta(t_2 - t_1) \\ &= i\hbar \sum_n \psi_n(\boldsymbol{x}_2)\psi_n^*(\boldsymbol{x}_1) e^{-iE_n(t_2-t_1)/\hbar}\delta(t_2 - t_1) \\ &= i\hbar \sum_n \psi_n(\boldsymbol{x}_2)\psi_n^*(\boldsymbol{x}_1)\delta(t_2 - t_1) \\ &= i\hbar\delta(\boldsymbol{x}_2 - \boldsymbol{x}_1)\delta(t_2 - t_1). \end{aligned} \tag{3.13}$$

The last step is obtained using the closure property (3.5). Although Eq. (3.13) is obtained assuming H to be time independent, it is true in the general case[2] when H is also time dependent.

$$\left\{ i\hbar \frac{\partial}{\partial t_2} - H(\boldsymbol{x}_2, t_2) \right\} K(\boldsymbol{x}_2 t_2; \boldsymbol{x}_1 t_1) = i\hbar\delta(\boldsymbol{x}_2 - \boldsymbol{x}_1)\delta(t_2 - t_1). \tag{3.14}$$

Equation (3.14) is the differential equation for the propagation kernel $K(\boldsymbol{x}_2 t_2; \boldsymbol{x}_1 t_1)$ and it is equivalent to the integral form (3.2).

Now, we are in a position to develop a perturbation theory. Suppose

$$H = H_0 + V,$$

[2]S. S. Schweber and H. A. Bethe, *Mesons and Fields, Vol. 1*, Row, Peterson & Co., New York (1956).

where V is a small perturbation. We know the solutions of the unperturbed Hamiltonian H_0. Then the propagation kernel K_0 for $V = 0$ corresponds to the free particle. This enables us to obtain an integral equation for K which is equivalent to the differential equation (3.14) with the boundary condition[3]

$$K(\mathbf{2},\mathbf{1}) = 0, \qquad t_2 < t_1.$$

and develop a perturbation theory for $K(\mathbf{2},\mathbf{1})$ in terms of $K_0(\mathbf{2},\mathbf{1})$. The integral equation for $K(\mathbf{2},\mathbf{1})$ is given below:

$$K(\mathbf{2},\mathbf{1}) = K_0(\mathbf{2},\mathbf{1}) - \frac{i}{\hbar c}\int K_0(\mathbf{2},\mathbf{3})V(\mathbf{3})K(\mathbf{3},\mathbf{1})d^4x_3, \tag{3.15}$$

where d^4x is the element of the four dimensional space-time volume[4].

$$d^4x = cdt\,dx\,dy\,dz = dx_0\,dx_1\,dx_2\,dx_3. \tag{3.16}$$

It can be easily verified that the integral equation (3.15) is equivalent to the differential equation

$$\left\{i\hbar\frac{\partial}{\partial t_2} - H_0(\mathbf{2})\right\}K(\mathbf{2},\mathbf{1}) = i\hbar\delta^4(\mathbf{2},\mathbf{1}) + V(\mathbf{2})K(\mathbf{2},\mathbf{1}). \tag{3.17}$$

Equation (3.17) is identical with Eq. (3.14). The integral equation (3.15) can be written as a perturbation series by repeated application of the expansion for K occurring on the right hand side of Eq. (3.15).

$$\begin{aligned} K(\mathbf{2},\mathbf{1}) &= K_0(\mathbf{2},\mathbf{1}) + \left(\frac{-i}{\hbar c}\right)\int K_0(\mathbf{2},\mathbf{3})V(\mathbf{3})K_0(\mathbf{3},\mathbf{1})d^4x_3 \\ &\quad + \left(\frac{-i}{\hbar c}\right)^2\int K_0(\mathbf{2},\mathbf{3})V(\mathbf{3})K_0(\mathbf{3},\mathbf{4})V(\mathbf{4})K_0(\mathbf{4},\mathbf{1})d^4x_3d^4x_4 \\ &\quad + \cdots . \end{aligned} \tag{3.18}$$

3.1.2 In relativistic Dirac theory

Consider the time-dependent Dirac equation in Feynman notation for a free particle[5] using natural units ($\hbar = c = 1$).

$$(i\nabla\!\!\!/ - m)\Psi(\boldsymbol{x},t) = 0. \tag{3.19}$$

[3]For brevity, we use the notation $\mathbf{1}$ for $(\boldsymbol{x}_1,t_1)$ and $\mathbf{2}$ for $(\boldsymbol{x}_2,t_2)$.

[4]The suffixes 1,2,3 are used to denote the space-time coordinates and also the x,y,z components but the reader can easily infer the correct meaning from the context they are used.

[5]The Dirac equation for a free particle in natural units can be written either as

$$(p\!\!\!/ - m)\Psi = 0 \qquad \text{or} \qquad (i\nabla\!\!\!/ - m)\Psi = 0,$$

Here $\Psi(\boldsymbol{x}_1, t_1)$ and $\Psi(\boldsymbol{x}_2, t_2)$ are four component wave functions and the propagation kernel $K(\boldsymbol{x}_2, t_2; \boldsymbol{x}_1, t_1)$ is a (4×4) matrix. The differential equation that the propagation kernel obeys is analogous to Eq. (3.14).

$$(i\not{\nabla} - m)K_0(\mathbf{2}, \mathbf{1}) = i\delta^4(\mathbf{2}, \mathbf{1}). \tag{3.20}$$

Expanding Ψ in terms of the complete set of eigenfunctions of (3.19), which correspond to both positive and negative energy eigenvalues, we obtain

$$K_0(\mathbf{2}, \mathbf{1}) = \begin{cases} \displaystyle\sum_{E_n>0} \psi_n(\boldsymbol{x}_2)\bar{\psi}_n(\boldsymbol{x}_1)e^{-iE_n(t_2-t_1)} \\ \qquad + \displaystyle\sum_{E_n<0} \psi_n(\boldsymbol{x}_2)\bar{\psi}_n(\boldsymbol{x}_1)e^{-iE_n(t_2-t_1)}, & t_2 > t_1, \\ 0, & t_2 < t_1. \end{cases} \tag{3.21}$$

For relativistic convenience, one deals with adjoint wave functions $\bar{\psi}$ instead of $\psi^\dagger$ in Eq. (3.21). The solution (3.21) is unsatisfactory since it corresponds to the one-electron theory instead of the correct hole theory. It admits the transition of the electron from the positive energy state to the negative energy state for $t_2 > t_1$. This is clearly unphysical since, according to the hole theory, all the negative energy states are full and hence not available for the electron after scattering. So, for $t_2 > t_1$, K_0 should contain only positive energy states. Each term $\psi_n(\boldsymbol{x}_2)\bar{\psi}_n(\boldsymbol{x}_1)e^{-iE_n(t_2-t_1)}$ is a solution of the homogeneous equation if it is used for all times t_2. So, it is permitted to subtract the sum over all the negative energy states for both $t_2 > t_1$ and $t_2 < t_1$. The resulting propagation kernel K_F, known as *Feynman's Propagation Kernel*, is given by

$$K_F(\mathbf{2}, \mathbf{1}) = \begin{cases} \displaystyle\sum_{E_n>0} \psi_n(\boldsymbol{x}_2)\bar{\psi}_n(\boldsymbol{x}_1)e^{-iE_n(t_2-t_1)}, & t_2 > t_1 \\ \displaystyle -\sum_{E_n<0} \psi_n(\boldsymbol{x}_2)\bar{\psi}_n(\boldsymbol{x}_1)e^{-iE_n(t_2-t_1)}, & t_2 < t_1. \end{cases} \tag{3.22}$$

Please note that $e^{-iE_n(t_2-t_1)} = e^{-i|E_n||t_2-t_1|}$, for both $t_2 > t_1$ and $t_2 < t_1$. The perturbation expansion (3.18) is still valid if K_0 is replaced by K_F. Only V is now a 4×4 matrix.

---

since $\not{p} = \gamma_\mu p_\mu = \gamma_0 E - \quad \cdot \boldsymbol{p}$, with $E = i\frac{\partial}{\partial t} = i\frac{\partial}{\partial x_0}$ and $\boldsymbol{p} = -i\boldsymbol{\nabla}$. Thus

$$\not{p} = i \quad \gamma_0 \frac{\partial}{\partial x_0} + \gamma_x \frac{\partial}{\partial x} + \gamma_y \frac{\partial}{\partial y} + \gamma_z \frac{\partial}{\partial z} \quad = i\not{\nabla}, \text{ where } \not{\nabla} = \gamma_0 \frac{\partial}{\partial x_0} + \gamma_x \frac{\partial}{\partial x} + \gamma_y \frac{\partial}{\partial y} + \gamma_z \frac{\partial}{\partial z}.$$

The choice of the propagation kernel (3.22) causes the positive energy components to be propagated forward in time while the negative energy components are propagated backwards in time. This is the most notable contribution of Feynman, which has resulted in the birth of Feynman's positron theory, according to which the electrons are propagated forward in time whereas the positrons correspond to negative energy electrons that are propagated backwards in time.

3.1.3 In momentum representation

The Feynman propagation kernel K_F obeys the differential equation[6]

$$(i\nabla\!\!\!/ - m)K_F(\mathbf{2},\mathbf{1}) = i\delta^4(\mathbf{2},\mathbf{1}). \tag{3.23}$$

If $S_+(\mathbf{p})$ is the Fourier transform of K_F, then

$$K_F(\mathbf{2},\mathbf{1}) = \int_{-\infty}^{+\infty} S_+(\mathbf{p}) e^{-i\mathbf{p}\cdot(\mathbf{x}_2-\mathbf{x}_1)} d^4p, \tag{3.24}$$

where

$$d^4p = dp_0 dp_x dp_y dp_z = dp_0 dp_1 dp_2 dp_3.$$

Denoting the four-dimensional delta function in the form of Fourier transform,

$$\delta^4(\mathbf{2},\mathbf{1}) = \frac{1}{(2\pi)^4} \int_{-\infty}^{+\infty} e^{-i\mathbf{p}\cdot(\mathbf{x}_2-\mathbf{x}_1)} d^4p, \tag{3.25}$$

Eq. (3.23) can be written as

$$(i\nabla\!\!\!/ - m) \int_{-\infty}^{+\infty} S_+(\mathbf{p}) e^{-i\mathbf{p}\cdot(\mathbf{x}_2-\mathbf{x}_1)} d^4p = \frac{i}{(2\pi)^4} \int_{-\infty}^{+\infty} e^{-i\mathbf{p}\cdot(\mathbf{x}_2-\mathbf{x}_1)} d^4p, \tag{3.26}$$

Equating the integrands in Eq. (3.26), we obtain

$$(p\!\!\!/ - m)S_+(\mathbf{p}) = \frac{i}{(2\pi)^4} \tag{3.27}$$

This yields the propagation kernel in the momentum representation.

$$S_+(\mathbf{p}) = \frac{i}{(2\pi)^4} \frac{1}{(p\!\!\!/ - m)} = \frac{i}{(2\pi)^4} \frac{(p\!\!\!/ + m)}{\mathbf{p}^2 - m^2}, \tag{3.28}$$

[6] We use bold italics $\boldsymbol{p}$, $\boldsymbol{x}$ to denote three-dimensional vectors and upright bold letters $\mathbf{p}$, $\mathbf{x}$ to denote four vectors.

since

$$(\not{p} + m)(\not{p} - m) = \not{p}^2 - m^2 \qquad \text{and} \qquad \not{p}^2 = p_0^2 - \boldsymbol{p}^2 = \mathbf{p}^2.$$

The foregoing discussion yields an expression for the propagation kernel $K_F(2,1)$.

$$K_F(\mathbf{2},\mathbf{1}) = \frac{i}{(2\pi)^4} \int \frac{1}{(\not{p} - m)} e^{-i\mathbf{p}\cdot(\mathbf{x}_2 - \mathbf{x}_1)} d^4p. \tag{3.29}$$

Equation (3.29) is obtained from Eqs. (3.24) and (3.28).

3.1.4 Interaction with electromagnetic field

Let us consider the interaction of an electron with an external electromagnetic field which is represented by an interaction potential

$$V = e\gamma_\mu A_\mu = e\not{A}. \tag{3.30}$$

The Dirac equation for an electron in an electromagnetic field is given by

$$(i\not{\nabla} - m)\Psi(\mathbf{x}) = e\not{A}\Psi(\mathbf{x}). \tag{3.31}$$

The propagation kernel K_F^A for the electron in an electromagnetic field satisfies the differential equation

$$(i\not{\nabla} - e\not{A} - m)K_F^A(2,1) = i\delta^4(2,1). \tag{3.32}$$

The propagation kernel has only positive energy components for $t_2 > t_1$ and only negative energy components for $t_2 < t_1$. It satisfies the integral equation

$$K_F^A(\mathbf{2},\mathbf{1}) \;=\; K_F(\mathbf{2},\mathbf{1}) - ie \int d^4x_3\, K_F(\mathbf{2},\mathbf{3})\not{A}(\mathbf{3})K_F^A(\mathbf{3},\mathbf{1}). \tag{3.33}$$

The perturbation expansion of this integral equation is given below.

$$\begin{aligned} K_F^A(\mathbf{2},\mathbf{1}) \;=\; & K_F(\mathbf{2},\mathbf{1}) - ie \int d^4x_3\, K_F(\mathbf{2},\mathbf{3})\not{A}(\mathbf{3})K_F(\mathbf{3},\mathbf{1}) \\ & + (-ie)^2 \int d^4x_3 \int d^4x_4\, K_F(\mathbf{2},\mathbf{3})\not{A}(\mathbf{3})K_F(\mathbf{3},\mathbf{4})\not{A}(\mathbf{4})K_F(\mathbf{4},\mathbf{1}) \\ & + \cdots . \end{aligned} \tag{3.34}$$

Let us now write down the Dirac wave function $\Psi(\mathbf{2})$ at the space-time point $\mathbf{2}$ in terms of the wave function $\Psi(\mathbf{1})$ at the space-time point $\mathbf{1}$.

$$\Psi(\mathbf{2}) = \int K_F^A(\mathbf{2},\mathbf{1})\beta\Psi(\mathbf{1})d^3x_1. \tag{3.35}$$

For the non-interacting case

$$\Psi_0(\mathbf{2}) = \int K_F(\mathbf{2},\mathbf{1})\beta\Psi_0(\mathbf{1})d^3x_1, \tag{3.36}$$

where Ψ_0 is a solution of the free Dirac equation.

3.2 The Transition Amplitude

3.2.1 First order matrix element

The transition amplitude for an electron to go from some positive energy state $\Psi_i(\boldsymbol{x}_1,t_1)$ at time t_1 to some other positive energy state $\Psi_f(\boldsymbol{x}_2,t_2)$ at time t_2 is given by

$$\begin{aligned}\mathscr{M} &= \int \Psi_f^\dagger(\boldsymbol{x}_2,t_2)K_F(\boldsymbol{x}_2t_2,\boldsymbol{x}_1t_1)\beta\Psi_i(\boldsymbol{x}_1,t_1)d^3x_1d^3x_2 \\ &= \int \bar{\Psi}_f(\mathbf{2})\beta K_F(\mathbf{2},\mathbf{1})\beta\Psi_i(\mathbf{1})d^3x_1d^3x_2. \end{aligned}\tag{3.37}$$

If an external potential acts between t_1 and t_2, then K_F is replaced by K_F^A and one can obtain the first order transition matrix element $\mathscr{M}_1$ using the perturbation series (3.34).

$$\mathscr{M}_1 = -ie\int\int\int \bar{\Psi}_f(2)\beta K_F(\mathbf{2},\mathbf{3})\not{A}(\mathbf{3})K_F(\mathbf{3},\mathbf{1})\beta\Psi_i(\mathbf{1})d^3x_1d^3x_2d^4x_3. \tag{3.38}$$

Since

$$\int K_F(\mathbf{3},\mathbf{1})\beta\Psi_i(\mathbf{1})d^3x_1 = \Psi_i(\mathbf{3}), \tag{3.39}$$

$$\int \bar{\Psi}_f(\mathbf{2})\beta K_F(\mathbf{2},\mathbf{3})d^3x_2 = \bar{\Psi}_f(\mathbf{3}), \tag{3.40}$$

the first order matrix element can be rewritten as

$$\mathscr{M}_1 = -ie\int \bar{\Psi}_f(\mathbf{3})\not{A}(\mathbf{3})\Psi_i(\mathbf{3})d^4x_3. \tag{3.41}$$

Let us choose the initial state to be the positive energy state of an electron with four-momentum $\mathbf{p}_1$ and the final state to be the positive energy state of the electron with four-momentum $\mathbf{p}_2$.

$$\Psi_i(\mathbf{x}) = u(\boldsymbol{p}_1)e^{-i\mathbf{p}_1\cdot\mathbf{x}}; \tag{3.42}$$

$$\bar{\Psi}_f(\mathbf{x}) = \bar{u}(\boldsymbol{p}_2)e^{i\mathbf{p}_2\cdot\mathbf{x}}. \tag{3.43}$$

If $a_\mu(\mathbf{k})$ is the Fourier transform of $A_\mu(\mathbf{x})$, then

$$A_\mu(\mathbf{x}) = \int a_\mu(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{x}}d^4k. \tag{3.44}$$

Substituting the results (3.42) – (3.44) in Eq. (3.41), we obtain

$$\mathscr{M}_1 = -ie\int \bar{u}(\boldsymbol{p}_2)\not{a}(\mathbf{k})u(\boldsymbol{p}_1)e^{-i(\mathbf{p}_2-\mathbf{k}-\mathbf{p}_1)\cdot\mathbf{x}}d^4xd^4k. \tag{3.45}$$

Integrating over d^4x, we get

$$\mathscr{M}_1 = -ie(2\pi)^4\int \bar{u}(\boldsymbol{p}_2)\not{a}(\mathbf{k})u(\boldsymbol{p}_1)\delta^4(\mathbf{p}_2-\mathbf{k}-\mathbf{p}_1)d^4k, \tag{3.46}$$

where $\delta^4(\mathbf{p}_2-\mathbf{k}-\mathbf{p}_1)$ denotes the energy-momentum conservation at the interaction point. The integration over d^4k is trivial and we finally obtain the first-order matrix element in a simple form.

$$\mathscr{M}_1 = -ie(2\pi)^4\bar{u}(\boldsymbol{p}_2)\not{a}(\mathbf{k})u(\boldsymbol{p}_1), \quad \text{with} \quad \mathbf{k} = \mathbf{p}_2 - \mathbf{p}_1. \tag{3.47}$$

3.2.2 Second order matrix element

In a similar way, we can obtain the second-order matrix element for the electron transition from the initial momentum state $\boldsymbol{p}_1$ to the final momentum state $\boldsymbol{p}_2$. Adopting the same method that we have used to obtain Eq. (3.41) for the first order matrix element $\mathscr{M}_1$, we obtain an expression for the second order matrix element.

$$\mathscr{M}_2 = (-ie)^2\int d^4x_3d^4x_4\bar{\Psi}_f(\mathbf{4})\not{A}(\mathbf{4})K_F(\mathbf{4},\mathbf{3})\not{A}(\mathbf{3})\Psi_i(\mathbf{3}). \tag{3.48}$$

Using expressions (3.42) – (3.44) and (3.29), and performing the integration over d^4x_3 and d^4x_4, we can go to the momentum representation to obtain

$$\mathscr{M}_2 = (-ie)^2\int d^4p\,d^4k_1\,d^4k_2\,(2\pi)^4\delta^4(\mathbf{p}_2-\mathbf{k}_2-\mathbf{p})\,\delta^4(\mathbf{p}-\mathbf{k}_1-\mathbf{p}_1)$$
$$\times\,\bar{u}(\boldsymbol{p}_2)\not{a}(\mathbf{k}_2)\frac{1}{\not{p}-m}\not{a}(\mathbf{k}_1)u(\boldsymbol{p}_1). \tag{3.49}$$

The integration over d^4p and d^4k_2 can be performed trivially because of the δ - functions and we are left with only one integration.

$$\mathscr{M}_2 = (-ie)^2(2\pi)^4 \int d^4k_1\, \bar{u}(\boldsymbol{p}_2) \not{\phi}(\mathbf{p}_2 - \mathbf{p}_1 - \mathbf{k}_1) \frac{1}{\not{p}_1 + \not{k}_1 - m} \not{\phi}(\mathbf{k}_1) u(\boldsymbol{p}_1). \tag{3.50}$$

Eqs. (3.47) and (3.50) have a simple physical interpretation. Eq. (3.47) describes an electron of momentum $\boldsymbol{p}_1$ interacting with an electromagnetic field, represented by $-ie\not{\phi}(\mathbf{k})$, picking up a four-momentum $\mathbf{k}$ and making a transition to the final state with momentum $\boldsymbol{p}_2$. In a similar way, Eq. (3.50) denotes an electron, that interacts with an electromagnetic field of four-momentum $\mathbf{k}_1$, propagates as a free particle with a kernel $\frac{1}{\not{p}_1 + \not{k}_1 - m}$ until it interacts once again with the electromagnetic field and emerges as an electron with momentum $\boldsymbol{p}_2$.

3.3 Feynman Diagrams

The propagation of electron interacting with an external field can be conveniently depicted by Feynman diagrams as shown in Fig. 3.1. (A) gives the Feynman diagrams in space-time representation and (B) gives the Feynman diagrams in momentum-energy representation. Both first order and second order transitions are considered for illustrative purpose.

The propagation of electron from space-time point (**1**) to space-time point (**2**) is usually called a world line. The world line need not always go forward in time but sometimes it can also go backwards in time.

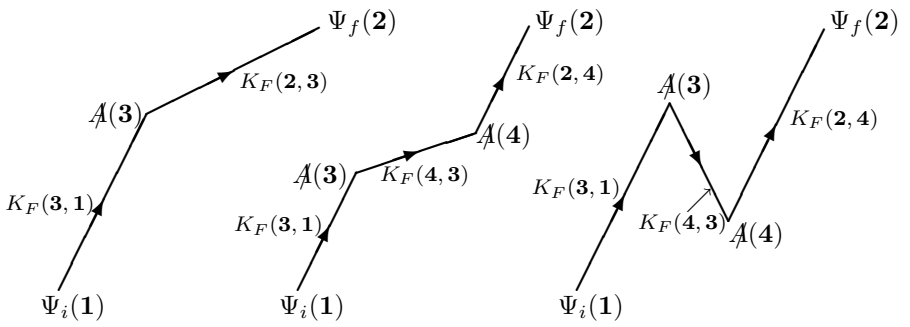
3.3.1 In space-time representation

Diagrams A(a) and A(b)

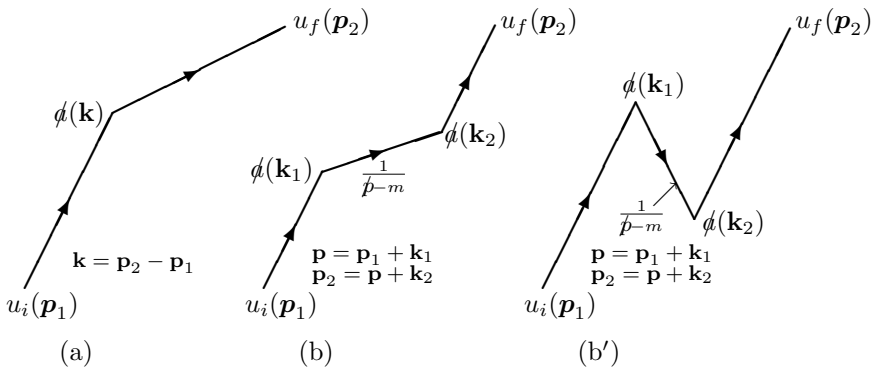
As we follow the world line from (**1**) to (**2**) in diagrams A(a) and A(b), time always increases. So in diagram A(a), the propagation kernels $K_F(\mathbf{3}, \mathbf{1})$ and $K_F(\mathbf{2}, \mathbf{3})$ consist of sums over positive energy states only as given by Eq. (3.22) for $t_2 > t_1$. Similar is the situation in case of diagram A(b) where all the propagation kernels $K_F(\mathbf{3}, \mathbf{1})$, $K_F(\mathbf{4}, \mathbf{3})$ and $K_F(\mathbf{2}, \mathbf{1})$ point towards the increasing time and hence consist of sums over positive energy states only. Diagram A(b) represents the conventional double scattering.

Diagram A(b′)

As we follow the world line, we find that we go backwards in time from **3** to **4** and hence the propagation kernel $K_F(\mathbf{4},\mathbf{3})$ consists of a sum of negative energy states only. Feynman interprets this as positron propagating from **4** to **3**. In the conventional view-point of increasing time, diagram A(b′) represents the following sequence of events. An electron-positron pair is created at space-time point **4**, the positron gets annihilated with the incoming electron at space-time point **3** and the electron propagates to **2** which is observed as the final state in scattering.



(A) In space-time representation



(B) In momentum-energy representation

Figure 3.1: Feynman diagrams (A) in space-time representation and (B) in momentum-energy representation representing the first order scattering (a) and the second order scattering (b and b′). The diagrams (b) and (b′) are equivalent. At each vertex, energy and momentum are conserved.

It is emphasized that both diagrams A(b) and A(b′) are included in

the single Eq. (3.48) since an integration is performed over all times t_3 and t_4. So, Eq. (3.48) is usually represented by a single diagram A(b) and the diagram A(b′) is considered identical to it, since it can be obtained from diagram A(b) when $t_4 < t_3$. This way of description satisfies the Dirac hole theory. A negative energy electron is lifted from the occupied negative energy sea to the electron final state **2** whereas the hole in the negative energy sea is filled by the incoming electron **1**. What is the observed effect that diagram A(b′) represents? An electron is scattered from state **1** to state **2** but at the same time a positive energy electron and a negative-sea electron is exchanged. The Pauli principle demands that the matrix element be given a negative sign. Indeed the Feynman propagation kernel $K_F(\mathbf{2}, \mathbf{1})$, given by Eq. (3.22), has a negative sign for $t_2 < t_1$ without explicitly using the Pauli principle. Thus the dynamics of Feynman's propagation kernel K_F include the negative sign as required by the Pauli principle. By this novel approach, Feynman is able to develop a theory which can describe quantum electrodynamical processes without recourse to the complications of old-fashioned hole theory with its intermediate states. Thus Feynman's approach greatly simplifies the calculational procedure and so, it has superseded all the older methods.

3.3.2 In momentum-energy representation

Figure 3.1(B) presents the Feynman diagrams of the afore-discussed events in momentum-energy representation. Diagram B(a) depicts the first order scattering of an electron by an external field. An electron with four-momentum $\mathbf{p}_1$ interacts with an external field and picks up a four-momentum $\mathbf{k}$ and reaches the final state with four-momentum $\mathbf{p}_2$. Diagrams B(b) and B(b′) depict the second order scattering. Both the diagrams describe an electron with an initial four-momentum $\mathbf{p}_1$, that is scattered by an eternal field, picking up a four-momentum $\mathbf{k}_1$ and travels as a virtual particle with a propagator $\frac{1}{\not{p}-m}$ and once again scattered, picking up a four-momentum $\mathbf{k}_2$ before reaching the final state with four-momentum $\mathbf{p}_2$. Although the topology of the two diagrams B(b) and B(b′) are slightly different due to the different ways of representing the intermediate state, both of them can be described mathematically by the same expression although they can be considered physically as arising from different physical processes as explained earlier in our discussion on space-time representation.

The momentum-energy representation of the events are of great importance to us since most of the calculations are done in this representation.

At each vertex, energy and momentum are conserved but at the intermediate states, the particle does not satisfy the required energy-momentum relation $E^2 = p^2 + m^2$ and hence the intermediate state particles are sometimes called virtual particles.

For calculational purpose, the momentum-energy representation is found to be more convenient and most of the quantum electrodynamic processes are studied by drawing Feynman diagrams in energy-momentum representation.

3.3.3 Basic vertices in QED

Feynman diagrams in Quantum Electrodynamics (QED) are made up of vertices which are points where three lines (electron, positron and photon lines) meet representing electromagnetic interaction with electron or positron. The basic vertices are: (a) absorption of a photon by an electron (b) emission of a photon by an electron (c) absorption of a photon by a positron (d) emission of a photon by a positron (e) electron-positron pair creation by a photon and (f) electron-positron pair annihilation into a photon. They are altogether six vertices as shown in Fig. 3.2. A single vertex is not a valid physical process since energy and momentum cannot be simultaneously conserved. So, a Feynman diagram for a valid physical process should consist of two or more such vertices and satisfy the following conditions:

1. Energy and momentum are conserved at each vertex.

2. The incoming and outgoing lines in a diagram are real particles that obey the relativistic energy-momentum relation $E^2 = p^2c^2 + m^2c^4$.

3. The internal lines (lines in the intermediate stages) are called the virtual particles that do not satisfy the energy-momentum relation but they form an integral part of the Feynman diagrams.

Any Feynman diagram consisting only of a single vertex cannot satisfy simultaneously the conditions 1 and 2 and hence cannot represent a physical process.

An electron is usually represented by a solid line with an arrow in the direction of increasing time, a positron by a solid line with an arrow opposite to the direction of increasing time and a photon by a wavy line.

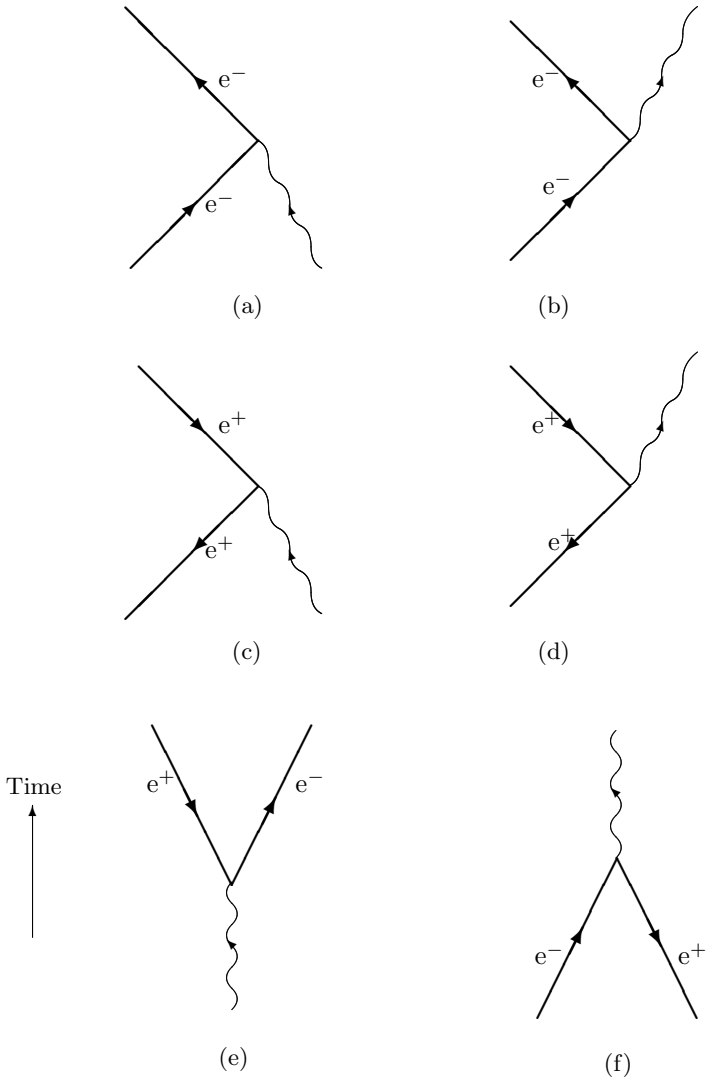


Figure 3.2: Feynman diagrams representing basic vertices in QED: (a) photon absorption by electron, (b) photon emission by electron, (c) photon absorption by positron, (d) photon emission by positron, (e) electron-positron pair creation, (f) electron-positron pair annihilation.

For a given process, draw all possible diagrams and write down the matrix element for each diagram and add them up. In other words, one has to compute the transition amplitude corresponding to each diagram and find the total amplitude merely by summing them. The total transition probability is obtained by finding the absolute square of the total amplitude. If the spins of the initial and final particles are not observed,

then a sum over the final spin states and an average over the initial spin states have to be taken.

Review Questions

3.1 Explain how you can study the time evolution of the state vector in Schrödinger equation by defining a propagation kernel $K_0(\mathbf{2},\mathbf{1})$. By introducing a perturbing potential V, show how the propagation kernel $K(\mathbf{2},\mathbf{1})$ with the perturbing potential can be expressed as an integral equation and obtain a perturbation series for $K(\mathbf{2},\mathbf{1})$ in terms of $K_0(\mathbf{2},\mathbf{1})$.

3.2 Obtain the propagation kernel for the wave function obeying the Dirac equation and explain how Feynman's modification gave a new insight that led to the development of Feynman's positron theory.

3.3 Obtain an expression for the transition amplitude for an electron to go from one positive energy state to another due to an interaction with electromagnetic field. Deduce the first order and second order matrix elements for such a transition both in space-time representation and momentum-energy representation.

3.4 Explain why the momentum-energy representation is preferable for the study of various processes in quantum electrodynamics.

Problems

3.1 It is said that a single vertex which corresponds to a first order matrix element is not a valid physical process. But the Coulomb scattering of an electron by a nucleus is represented by a single vertex and calculated by using the first order matrix element. Explain the contradiction.

3.2 Draw the Feynman diagrams to describe the Compton scattering. Show that there are two possible Feynman diagrams which contribute to the process. Write down the matrix element corresponding to each diagram.

3.3 Show that there are two Feynman diagrams that contribute to the electron-electron scattering process in the lowest order. Obtain the matrix element for the two diagrams.

Solutions to Problems

3.1 The Coulomb scattering of an electron by a nucleus is represented by the Feynman diagram as given in Fig. 4.1. Please note that the wavy line in figure denotes a virtual photon and not a real photon. A single vertex does not represent a physical process only if all the particles at a single vertex are real. So, there is no contradiction and Fig. 4.1 represents a valid physical process and so the cross section can be calculated using the first order matrix element (3.47).

3.2 The two Feynman diagrams for the Compton scattering are given in Fig. 4.3. It is a second order process and the matrix elements can be written down using Eq. (3.50). For details, refer to Chapter 4.

3.3 The problem of electron-electron scattering is treated in Sec. 4.3. along with Feynman diagrams. The reader is referred to that section for writing down the Feynman amplitudes for the process.

Chapter 4

Quantum Electrodynamics

Quantum Electrodynamics (QED) deals with processes involving interaction of charged particles with electromagnetic field. In this chapter, we shall deal with lowest order Feynman diagrams that describe the processes involving electrons and photons. If we go to higher orders, we are faced with problems of infinities and we shall briefly discuss how to tame them by renormalization techniques in the next chapter.

There are several processes, of which we shall choose the following for discussion: (1) Rutherford scattering (2) Compton scattering (3) Electron-electron scattering (4) Electron-positron scattering (5) Electron-positron pair annihilation (6) Bremsstrahlung (7) Electron-positron pair production and (8) Muon pair production in electron-positron collision. As Schwinger once remarked, Feynman has indeed taken the quantum elecrodynamical calculations "to the masses".

4.1 Rutherford Scattering

Consider the high energy scattering of an electron by the Coulomb potential of a nucleus of charge Ze as depicted in Fig. 4.1. The potential scattering causes only a three-momentum transfer and no energy transfer. The matrix element for the process can be written down[1] using Eq. (3.47).

$$\mathscr{M}_1 = -ie(2\pi)^4 \bar{\psi}(\boldsymbol{p}_f) \not{\phi}(\mathbf{q}) \psi(\boldsymbol{p}_i). \tag{4.1}$$

[1]We denote the four-vectors by upright bold letters $(\mathbf{p}, \mathbf{x})$, the three-vectors by italic bold letters $(\boldsymbol{p}, \boldsymbol{x})$ and the scalars by ordinary italics (p, x).

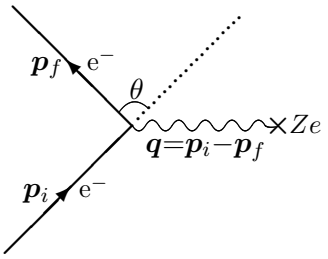


Figure 4.1: Coulomb scattering of electron by a nucleus of charge Ze. $\boldsymbol{p}_i$ and $\boldsymbol{p}_f$ denote the momenta of the incoming and scattered electron and $\boldsymbol{q}$, the three-momentum transfer to the electron. Since there is no energy transfer $|\boldsymbol{p}_i| = |\boldsymbol{p}_f| = p$ and $q^2 = 4p^2 \sin^2 \frac{\theta}{2}$.

where $\mathbf{q} = \mathbf{p}_f - \mathbf{p}_i$ is the four-momentum transfer $(q_0, \boldsymbol{q})$ to the electron in the Coulomb interaction $V = Ze/r$ with a nucleus of charge Ze, r being the distance measured from the centre of the nucleus. From Eq. (3.44), we obtain[2]

$$\begin{aligned} \not{A}(\mathbf{q}) &= \frac{1}{(2\pi)^4} \int e^{i\mathbf{q}\cdot\mathbf{x}} \gamma_\mu A_\mu(\mathbf{x}) d^4x \\ &= \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} e^{iq_0x_0} dx_0 \int e^{-i\boldsymbol{q}\cdot\boldsymbol{x}} \gamma_0 V(\boldsymbol{x}) d^3x, \end{aligned} \tag{4.2}$$

since $A_0(\mathbf{x}) = V(\boldsymbol{x}) = V(\boldsymbol{r}) = \frac{Ze}{r}$ and $A_k(\mathbf{x}) = 0,\ k = 1, 2, 3$ for a Coulomb interaction. The integrals in Eq. (4.2) can be evaluated using spherical coordinates and the results are given below.

$$\int_{-\infty}^{\infty} e^{iq_0x_0} dx_0 = 2\pi\delta(q_0). \tag{4.3}$$

$$\int e^{-i\boldsymbol{q}\cdot\boldsymbol{x}} V(\boldsymbol{x}) d^3x = Ze \int e^{-i\boldsymbol{q}\cdot\boldsymbol{r}} \frac{1}{r} d^3r = \frac{4\pi Ze}{q^2}. \tag{4.4}$$

Substituting Eqs. (4.3) and (4.4) in Eq. (4.2), we get

$$\not{A}(\mathbf{q}) = \frac{1}{(2\pi)^3} \frac{4\pi Ze}{q^2} \gamma_0 \delta(q_0). \tag{4.5}$$

[2]The convention of summation over repeated indices is used. Greek letters μ, ν, λ, ρ are used as suffixes to denote the components of four-vectors and Latin letters i, j, k, l are used as suffixes to denote the components of three-vectors.

Substituting this $\not{\phi}(\mathbf{q})$ in Eq. (4.1) and taking its absolute square, we obtain the transition probability.

$$|\mathscr{M}_1|^2 = (2\pi)^2 \left(\frac{4\pi Z e^2}{q^2}\right)^2 |\bar{\psi}(\boldsymbol{p}_f)\gamma_0\psi(\boldsymbol{p}_i)|^2[\delta(q_0)]^2, \tag{4.6}$$

where the square of the delta function requires some interpretation. Writing

$$[\delta(q_0)]^2 = \delta(q_0)\,\delta(0), \tag{4.7}$$

the first delta function can be interpreted to mean the conservation of energy and the second delta function $\delta(0)$ can be interpreted to denote the interaction time as shown below.

$$\delta(0) = \lim_{T\to\infty}\lim_{x\to 0}\frac{1}{2\pi}\int_{-T/2}^{+T/2} e^{ixt}dt = \lim_{T\to\infty}\frac{T}{2\pi}, \tag{4.8}$$

where T is interpreted as the time during which interaction takes place. From Eqs. (4.6) - (4.8), we obtain the transition probability per unit time to a group of final states in the continuum with energy E_f.

$$w_{i\to f} = 2\pi\left(\frac{4\pi Z e^2}{q^2}\right)^2 |\bar{\psi}(\boldsymbol{p}_f)\gamma_0\psi(\boldsymbol{p}_i)|^2\delta(E_i - E_f)\rho_f, \tag{4.9}$$

where ρ_f denotes the density of final states.

If we are considering an initial beam of unpolarized electron and not interested in the final spin states of the electron, then a sum over the final spin states and an average over the initial spin states have to be taken. For this, we use Eq. (2.68).

$$\overline{\sum}|\bar{\psi}(\boldsymbol{p}_f)\gamma_0\psi(\boldsymbol{p}_i)|^2 = \frac{1}{2}\mathrm{Tr}\left(\gamma_0(\not{p}_i + m)\,\tilde{\gamma_0}\,(\not{p}_f + m)\right), \tag{4.10}$$

where $\tilde{\gamma_0} = \gamma_0\gamma_0^\dagger\gamma_0 = \gamma_0$. and $\not{p} = \gamma_0 E - \boldsymbol{\gamma}\cdot\boldsymbol{p}$. Since the trace[3] of a

[3]The reader is advised to refer to Sec. 2.2 (Algebra of γ matrices) for details of evaluation of traces.

product of an odd number of γ matrices vanishes, we get[4]

$$\begin{aligned}\overline{\sum}|\bar{\psi}(\boldsymbol{p}_f)\gamma_0\psi(\boldsymbol{p}_i)|^2 &= \frac{1}{2}\mathrm{Tr}\left(\gamma_0 \underbrace{\not{p}_i \gamma_0}\, \not{p}_f + m^2\right)\\ &= \frac{1}{2}\mathrm{Tr}\left((-\not{p}_i\gamma_0 + 2E_i)\gamma_0\not{p}_f + m^2\right)\\ &= \frac{1}{2}\mathrm{Tr}\left(-\not{p}_i\not{p}_f + 2E_i\gamma_0\not{p}_f + m^2\right)\\ &= -2\,\mathbf{p}_i\cdot\mathbf{p}_f + 4E_iE_f + 2m^2. \qquad (4.11)\end{aligned}$$

Expanding the scalar product of the four-vectors $\mathbf{p}_i\cdot\mathbf{p}_f = E_iE_f - \boldsymbol{p}_i\cdot\boldsymbol{p}_f$ and rearranging, we get

$$\overline{\sum}|\bar{\psi}(\boldsymbol{p}_f)\gamma_0\psi(\boldsymbol{p}_i)|^2 = 2E_iE_f + 2p_ip_f\cos\theta + 2m^2, \qquad (4.12)$$

where θ denotes the angle of scattering. In the potential scattering that we are considering, there is no energy transfer but there is only a momentum transfer. $E_i = E_f = E$ and $|\boldsymbol{p}_i| = |\boldsymbol{p}_f| = p$. Hence

$$\begin{aligned}\overline{\sum}|\bar{\psi}(\boldsymbol{p}_f)\gamma_0\psi(\boldsymbol{p}_i)|^2 &= 2E^2 + 2p^2\cos\theta + 2m^2\\ &= 4E^2 - 2p^2(1-\cos\theta),\ \text{since } m^2 = E^2 - p^2\\ &= 4E^2 - 4p^2\sin^2(\theta/2)\\ &= 4E^2(1 - v^2\sin^2(\theta/2)), \qquad (4.13)\end{aligned}$$

since[5] $v = p/E$ is the velocity of the incident electron.

The density of final states ρ_f is given by

$$\rho_f = \frac{dn}{dE_f} = \frac{d^3p_f}{(2\pi)^3 dE_f} = \frac{p_f^2 dp_f d\Omega}{(2\pi)^3 dE_f} = \frac{p_f E_f d\Omega}{(2\pi)^3}. \qquad (4.14)$$

The last step is obtained using the relation $p_f dp_f = E_f dE_f$.

Substituting (4.13) and (4.14) in Eq. (4.9) and introducing the normalization factor $\sqrt{2E}$ for each fermion wave function such that $N = 2E$ and remembering that $E_f = E$ and $p_f = p$, we get the transition probability per unit time (the transition rate).

$$w_{i\to f} = \frac{2\pi}{N^2}\left(\frac{4\pi Ze^2}{q^2}\right)^2 4E^2\left(1 - v^2\sin^2\frac{\theta}{2}\right)\frac{pE}{(2\pi)^3}. \qquad (4.15)$$

[4] Underbraces are used to highlight the factors that are considered in the step-by-step calculation for the benefit of the reader.

[5] $E^2 = p^2 + m^2$. Differentiating, we get $2EdE = 2pdp$. Therefore $v = dE/dp = p/E$.

Since the incident wave is normalized to 1 particle per unit volume, the incident flux is v.

$$\text{Cross Section } = \frac{\text{Transition Rate}}{\text{Incident Flux}}$$

Thus

$$\frac{d\sigma}{d\Omega} = \frac{2\pi}{v}\left(\frac{4\pi Ze^2}{q^2}\right)^2\left(1 - v^2\sin^2\frac{\theta}{2}\right)\frac{pE}{(2\pi)^3}. \tag{4.16}$$

Substituting $q^2 = 4p^2\sin^2(\theta/2)$ in Eq. (4.16) and simplifying, we get

$$\frac{d\sigma}{d\Omega} = \frac{1}{v^2}\frac{Z^2e^4}{4p^2\sin^4(\theta/2)}\left(1 - v^2\sin^2\frac{\theta}{2}\right). \tag{4.17}$$

Equation (4.17) is the Mott scattering cross section. In the non-relativistic limit, since $v = \frac{p}{E} << 1$, the cross section (4.17) reduces to the Rutherford scattering cross section.

$$\frac{d\sigma}{d\Omega} = \frac{1}{v^2}\frac{Z^2e^4}{4p^2\sin^4(\theta/2)}. \tag{4.18}$$

The relativistic effect is to introduce the additional factor $(1-v^2\sin^2(\theta/2))$ to the Rutherford scattering formula.

4.1.1 Fermi's golden rule

Above, we have deduced the cross section (4.17) from first principles. But it is found convenient to use Fermi's golden rule to obtain the transition probability per second (transition rate) and differential cross section.

$$\text{Transition rate} = 2\pi(\Pi N)^{-1}|\overline{\mathcal{M}}|^2\rho_f; \tag{4.19}$$

$$d\sigma = \frac{\text{Transition rate}}{\text{Incident flux}} = \frac{2\pi}{v}(\Pi N)^{-1}|\overline{\mathcal{M}}|^2\rho_f; \tag{4.20}$$

where $(\Pi N)^{-1}$ denotes Feynman's normalization factors, $|\overline{\mathcal{M}}|^2$ stands for the square of the transition matrix element, obtained after summing over the final spin states and averaging over the initial spin states and ρ_f is

the density of final states[6]. In the present case,

$$(\Pi N)^{-1} = \frac{1}{4E^2}, \tag{4.21}$$

$$|\overline{\mathcal{M}}|^2 = \frac{1}{2}\sum_{\text{spins}} |\langle u_f|\mathcal{O}|u_i\rangle|^2, \quad \text{with } \mathcal{O} = \frac{4\pi z e^2}{q^2}\gamma_0, \tag{4.22}$$

$$\rho_f = \frac{pE}{(2\pi)^3} d\Omega. \tag{4.23}$$

Substituting (4.21) - (4.23) into Eq. (4.20), we obtain the Mott differential cross section (4.17).

In the consideration of the other quantum electrodynamic processes, discussed below in this chapter, we shall use Fermi's golden rule for calculating the differential cross sections.

4.2 Compton Scattering

The scattering of a photon by an electron is known as *Compton scattering*, for which the relativistic cross section was first deduced by Klein and Nishina[7]. The potential corresponding to the absorption of a photon is given by

$$A_\mu = \left(\frac{4\pi e^2}{2\omega}\right)^{1/2} \epsilon_\mu \, e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{4.24}$$

For the emission of photon, one has to take A_μ^*, the complex conjugate of (4.24). The potential A_μ in (4.24) is normalized to 1 photon per unit volume which is not relativistically invariant. For relativistic invariance, the potential should be normalized to 2ω photons per unit volume. So, the relativistically invariant potential for the absorption of a photon can be written as

$$A_\mu = \left(4\pi e^2\right)^{1/2} \epsilon_\mu \, e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{4.25}$$

To obtain the correct transition probability in a given coordinate system, it is necessary to reinsert a factor $(2\omega)^{-1}$ for each photon in the initial and final states. This becomes a part of the normalization factor ΠN which contains a similar factor for each electron in the initial and final states.

[6]The reader is referred to Appendix B for a comprehensive discussion on density of final states.

[7]O. Klein and Y. Nishina, *Z. Physik*, **52**, 853 (1929).

In momentum representation, the amplitude to absorb (emit) a photon of polarization ϵ_μ is

$$-i(4\pi e^2)^{1/2}\not{\epsilon}\,.$$

The polarization vector $\boldsymbol{\epsilon}$ is perpendicular to the propagation vector $\boldsymbol{k}$.

$$\boldsymbol{\epsilon}\cdot\boldsymbol{k}=0.$$

4.2.1 Kinematics

As a first approximation, consider the electrons to be free and choose the laboratory system, in which the initial electron is at rest. A photon of four momentum $\mathbf{k}_1$ is incident along the x-axis on an electron at rest and the scattered photon makes an angle θ with the incident direction and carries a four-momentum $\mathbf{k}_2$ as shown in Fig. 4.2. The electron recoils at an angle ϕ.

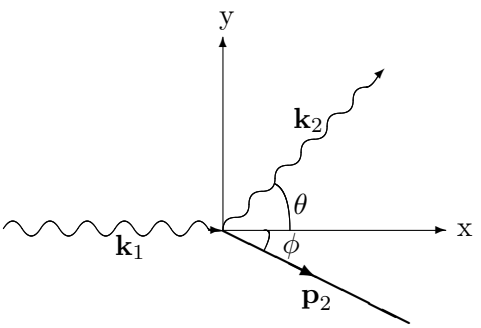


Figure 4.2: Compton scattering in the rest frame of the initial electron.

We shall represent the incoming and outgoing photons by the potentials

$$A_{1\mu}=\epsilon_{1\mu}e^{-i\mathbf{k}_1\cdot\mathbf{x}},\qquad A_{2\mu}=\epsilon_{2\mu}e^{-i\mathbf{k}_2\cdot\mathbf{x}},\tag{4.26}$$

and the initial and final electron states by the wave functions

$$\psi_1=u_1e^{-i\mathbf{p}_1\cdot\mathbf{x}},\qquad \psi_2=u_2e^{-i\mathbf{p}_2\cdot\mathbf{x}}.\tag{4.27}$$

The electron spinors obey the Dirac equation

$$\not{p}_1u_1=mu_1,\qquad \not{p}_2u_2=mu_2,\tag{4.28}$$

such that

$$\mathbf{p}_1\cdot\mathbf{p}_1=m^2,\qquad \mathbf{p}_2\cdot\mathbf{p}_2=m^2.\tag{4.29}$$

For a photon, the energy and momentum are both equal to frequency in natural units ($\hbar = c = 1$) such that

$$\mathbf{k}_1 \cdot \mathbf{k}_1 = 0, \qquad \mathbf{k}_2 \cdot \mathbf{k}_2 = 0. \tag{4.30}$$

The polarization four-vector has no time component but its space components are normal to its momentum vector. Hence

$$\boldsymbol{\varepsilon}_1 \cdot \mathbf{k}_1 = 0, \qquad \boldsymbol{\varepsilon}_2 \cdot \mathbf{k}_2 = 0. \tag{4.31}$$

Conservation of energy and momentum is given by

$$\mathbf{p}_1 + \mathbf{k}_1 = \mathbf{p}_2 + \mathbf{k}_2 \quad \text{or} \quad \not{p}_1 + \not{k}_1 = \not{p}_2 + \not{k}_2. \tag{4.32}$$

One of the angles, either θ or ϕ, is sufficient to determine the remaining quantities. Choosing the rest system of the initial electron, we get

$$\begin{aligned} \not{p}_2 &= \not{p}_1 + \not{k}_1 - \not{k}_2; \\ \not{p}_2^2 &= (\not{p}_1 + \not{k}_1 - \not{k}_2)(\not{p}_1 + \not{k}_1 - \not{k}_2) \\ &= \mathbf{p}_1^2 + \mathbf{k}_1^2 + \mathbf{k}_2^2 + 2\mathbf{p}_1 \cdot \mathbf{k}_1 - 2\mathbf{p}_1 \cdot \mathbf{k}_2 - 2\mathbf{k}_1 \cdot \mathbf{k}_2 \\ &= m^2 + 2m\omega_1 - 2m\omega_2 - 2\omega_1\omega_2(1 - \cos\theta). \end{aligned} \tag{4.33}$$

Since $\not{p}_2^2$ is also equal to

$$\not{p}_2^2 = \mathbf{p}_2 \cdot \mathbf{p}_2 = m^2, \tag{4.34}$$

we obtain the following relation from Eqs. (4.33) and (4.34).

$$\begin{aligned} m(\omega_1 - \omega_2) &= \omega_1\omega_2(1 - \cos\theta); \\ \frac{m}{\omega_2} - \frac{m}{\omega_1} &= 1 - \cos\theta. \end{aligned} \tag{4.35}$$

This is the well-known formula for the Compton shift in frequency or wavelength.

4.2.2 Transition matrix element

Compton scattering is a process in which an electron of four-momentum $\mathbf{p}_1$ and photon of four-momentum $\mathbf{k}_1$ in the initial state make a transition to a final state of electron of four-momentum $\mathbf{p}_2$ and photon of four-momentum $\mathbf{k}_2$. In the lowest order, it can be envisaged as a second-order process in which an electron first absorbs a photon of four-momentum

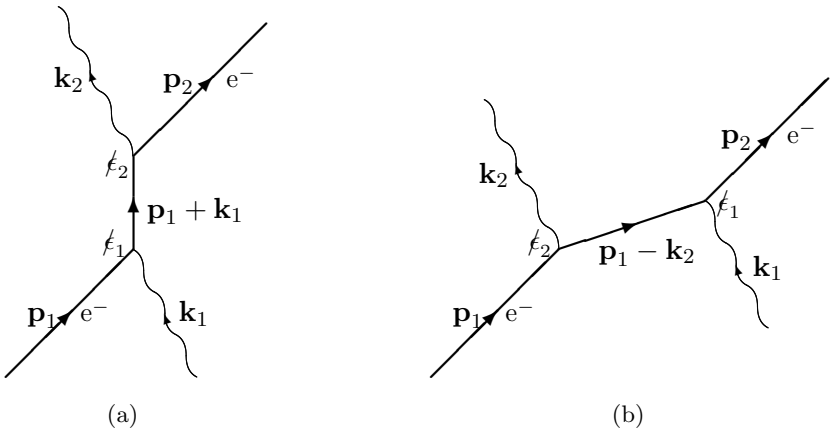


Figure 4.3: Lowest order Feynman diagrams for Compton scattering.

$\mathbf{k}_1$ and then emits a photon of four-momentum $\mathbf{k}_2$. This process is indistinguishable from a process in which the electron first emits a photon of four-momentum $\mathbf{k}_2$ and then absorbs a photon of four-momentum $\mathbf{k}_1$. These two modes can be represented by Feynman diagrams (a) and (b) as shown in Fig. 4.3.

Using the Feynman rules, the matrix elements for the diagrams (a) and (b) can be written as

$$\mathcal{M}_{\rm a} = -i4\pi e^2 \left(\bar{u}_2 \not{\epsilon}_2 \frac{1}{\not{p}_1 + \not{k}_1 - m} \not{\epsilon}_1 u_1 \right), \tag{4.36}$$

$$\mathcal{M}_{\rm b} = -i4\pi e^2 \left(\bar{u}_2 \not{\epsilon}_1 \frac{1}{\not{p}_1 - \not{k}_2 - m} \not{\epsilon}_2 u_1 \right). \tag{4.37}$$

Taking the sum of these two contributions, we get

$$\mathcal{M} = \mathcal{M}_{\rm a} + \mathcal{M}_{\rm b} = -i4\pi e^2 \left\{ \bar{u}_2 (\mathcal{O}_{\rm a} + \mathcal{O}_{\rm b}) u_1 \right\}, \tag{4.38}$$

with

$$\mathcal{O}_{\rm a} = \not{\epsilon}_2 \frac{1}{\not{p}_1 + \not{k}_1 - m} \not{\epsilon}_1 = \frac{\not{\epsilon}_2 (\not{p}_1 + \not{k}_1 + m) \not{\epsilon}_1}{(\not{p}_1 + \not{k}_1)^2 - m^2}, \tag{4.39}$$

$$\mathcal{O}_{\rm b} = \not{\epsilon}_1 \frac{1}{\not{p}_1 - \not{k}_2 - m} \not{\epsilon}_2 = \frac{\not{\epsilon}_1 (\not{p}_1 - \not{k}_2 + m) \not{\epsilon}_2}{(\not{p}_1 - \not{k}_2)^2 - m^2}. \tag{4.40}$$

Here, we make an interesting observation. The matrix element $\mathcal{M}_{\rm a}$ goes over into $\mathcal{M}_{\rm b}$ by making the substitution

$$\mathbf{k}_1 \longrightarrow -\mathbf{k}_2 \qquad \boldsymbol{\varepsilon}_1 \longrightarrow \boldsymbol{\varepsilon}_2. \tag{4.41}$$

and the total matrix element (the total transition amplitude) is invariant by the substitution (4.41). Such symmetries are known as *crossing symmetries.*

We can simplify the expressions (4.39) and (4.40) by choosing a coordinate system in which the initial electron is at rest. The denominators now simplify to

$$\begin{aligned}(\not{p}_1 + \not{k}_1)^2 - m^2 &= \not{p}_1^2 + \not{k}_1^2 + 2\not{p}_1\not{k}_1 - m^2 \\ &= \mathbf{p}_1^2 + \mathbf{k}_1^2 + 2\mathbf{p}_1 \cdot \mathbf{k}_1 - m^2 = 2m\omega_1, \qquad (4.42)\\ (\not{p}_1 - \not{k}_2)^2 - m^2 &= \not{p}_1^2 + \not{k}_2^2 - 2\not{p}_1\not{k}_2 - m^2 \\ &= \mathbf{p}_1^2 + \mathbf{k}_2^2 - 2\mathbf{p}_1 \cdot \mathbf{k}_2 - m^2 = -2m\omega_2, \qquad (4.43)\end{aligned}$$

since

$$\mathbf{p}_1^2 = m^2, \quad \mathbf{k}_1^2 = 0, \quad 2\mathbf{p}_1 \cdot \mathbf{k}_1 = 2m\omega_1, \quad \mathbf{k}_2^2 = 0, \quad 2\mathbf{p}_1 \cdot \mathbf{k}_2 = 2m\omega_2.$$

The numerators also can be similarly simplified.

$$\begin{aligned}\not{\epsilon}_2\not{p}_1\not{\epsilon}_1 &= -\not{\epsilon}_2\not{\epsilon}_1\not{p}_1 + 2\mathbf{p}_1 \cdot \boldsymbol{\varepsilon}_1 = -\not{\epsilon}_2\not{\epsilon}_1\not{p}_1, \qquad (4.44)\\ \not{\epsilon}_1\not{p}_1\not{\epsilon}_2 &= -\not{\epsilon}_1\not{\epsilon}_2\not{p}_1 + 2\mathbf{p}_1 \cdot \boldsymbol{\varepsilon}_2 = -\not{\epsilon}_1\not{\epsilon}_2\not{p}_1, \qquad (4.45)\end{aligned}$$

since $\mathbf{p}_1$ has only a time component whereas $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ have only space components such that $\mathbf{p}_1 \cdot \boldsymbol{\varepsilon}_1 = 0$ and $\mathbf{p}_1 \cdot \boldsymbol{\varepsilon}_2 = 0$. Further when taken between the electron spinors, they yield

$$\begin{aligned}\bar{u}_2\not{\epsilon}_2\not{p}_1\not{\epsilon}_1 u_1 &= -\bar{u}_2\not{\epsilon}_2\not{\epsilon}_1\not{p}_1 u_1 = -m\bar{u}_2\not{\epsilon}_2\not{\epsilon}_1 u_1, \qquad (4.46)\\ \bar{u}_2\not{\epsilon}_1\not{p}_1\not{\epsilon}_2 u_1 &= -\bar{u}_2\not{\epsilon}_1\not{\epsilon}_2\not{p}_1 u_1 = -m\bar{u}_2\not{\epsilon}_1\not{\epsilon}_2 u_1, \qquad (4.47)\end{aligned}$$

since $\not{p}_1 u_1 = m u_1$. Substituting (4.42) - (4.47) into Eqs. (4.39) and (4.40) and simplifying, we get

$$\begin{aligned}\tilde{u_2}\mathcal{O}_{\mathrm{a}} u_1 &= \frac{\bar{u}_2\not{\epsilon}_2\not{k}_1\not{\epsilon}_1 u_1}{2m\omega_1} = -\frac{\bar{u}_2\not{\epsilon}_2\not{\epsilon}_1\not{k}_1 u_1}{2m\omega_1}; \qquad (4.48)\\ \bar{u}_2\mathcal{O}_{\mathrm{b}} u_1 &= \frac{\bar{u}_2\not{\epsilon}_1\not{k}_2\not{\epsilon}_2 u_1}{2m\omega_2} = -\frac{\bar{u}_2\not{\epsilon}_1\not{\epsilon}_2\not{k}_2 u_1}{2m\omega_2}. \qquad (4.49)\end{aligned}$$

The last step in Eqs. (4.48) is obtained by using the commutation relation $\not{k}_1\not{\epsilon}_1 = -\not{\epsilon}_1\not{k}_1 + 2\mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_1$. The scalar product of four vectors vanishes since $\boldsymbol{\varepsilon}_1$ has no time component and the polarization three vector is normal to the photon three-momentum. A similar argument leads also to the last step in Eq. (4.49).

4.2.3 Transition probability

The absolute square of the transition matrix element $\mathcal{M}$ yields the transition probability. If the spin states of the electron in the initial and final states are not observed, then a sum over the final spin states and an average over the initial spin states have to be done. This is done by using the trace techniques discussed earlier.

$$|\overline{\mathcal{M}}|^2 = 16\pi^2 e^4 \overline{\sum} |\bar{u}_2 \mathcal{O} u_1|^2, \tag{4.50}$$

where[8]

$$\begin{aligned} \mathcal{O} &= \mathcal{O}_{\rm a} + \mathcal{O}_{\rm b} \\ &= \frac{\not\epsilon_2 \not\epsilon_1 \not k_1}{2m\omega_1} + \frac{\not\epsilon_1 \not\epsilon_2 \not k_2}{2m\omega_2}. \end{aligned} \tag{4.51}$$

Using the trace techniques, we get

$$\begin{aligned} \overline{\sum} |\bar{u}_2 \mathcal{O} u_1|^2 &= \frac{1}{2} {\rm Tr} \left[\mathcal{O}(\not p_1 + m) \tilde{\mathcal{O}} (\not p_2 + m) \right] \\ &= \frac{1}{2} {\rm Tr} \left[(\mathcal{O}_{\rm a} + \mathcal{O}_{\rm b})(\not p_1 + m)(\tilde{\mathcal{O}}_{\rm a} + \tilde{\mathcal{O}}_{\rm b})(\not p_2 + m) \right] \\ &= \frac{1}{2} \left[T_{\rm aa} + T_{\rm ab} + T_{\rm ba} + T_{\rm bb} \right], \end{aligned} \tag{4.52}$$

where

$$\tilde{\mathcal{O}} = \gamma_0 \mathcal{O}^\dagger \gamma_0, \quad \tilde{\mathcal{O}}_{\rm a} = \gamma_0 \mathcal{O}_{\rm a}{}^\dagger \gamma_0, \quad \tilde{\mathcal{O}}_{\rm b} = \gamma_0 \mathcal{O}_{\rm b}^\dagger \gamma_0, \tag{4.53}$$

and

$$T_{\rm aa} = {\rm Tr} \left[\mathcal{O}_{\rm a}(\not p_1 + m) \tilde{\mathcal{O}}_{\rm a} (\not p_2 + m) \right]; \tag{4.54}$$

$$T_{\rm ab} = {\rm Tr} \left[\mathcal{O}_{\rm a}(\not p_1 + m) \tilde{\mathcal{O}}_{\rm b} (\not p_2 + m) \right]; \tag{4.55}$$

$$T_{\rm ba} = {\rm Tr} \left[\mathcal{O}_{\rm b}(\not p_1 + m) \tilde{\mathcal{O}}_{\rm a} (\not p_2 + m) \right]; \tag{4.56}$$

$$T_{\rm bb} = {\rm Tr} \left[\mathcal{O}_{\rm b}(\not p_1 + m) \tilde{\mathcal{O}}_{\rm b} (\not p_2 + m) \right]. \tag{4.57}$$

[8] In Eq. (4.51), $-\mathcal{O}$ is replaced by $\mathcal{O}$, since the negative sign is of no consequence in the calculation of transition probability and hence omitted.

Writing explicitly the operators $\mathcal{O}_{\rm a}$, $\mathcal{O}_{\rm b}$, $\tilde{\mathcal{O}}_{\rm a}$ and $\tilde{\mathcal{O}}_{\rm b}$,

$$\mathcal{O}_{\rm a} = \frac{1}{2m\omega_1}\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1, \tag{4.58}$$

$$\mathcal{O}_{\rm b} = \frac{1}{2m\omega_2}\not{\varepsilon}_1\not{\varepsilon}_2\not{k}_2, \tag{4.59}$$

$$\tilde{\mathcal{O}}_{\rm a} = \frac{1}{2m\omega_1}\gamma_0(\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1)^\dagger\gamma_0 = \frac{1}{2m\omega_1}(\not{k}_1\not{\varepsilon}_1\not{\varepsilon}_2), \tag{4.60}$$

$$\tilde{\mathcal{O}}_{\rm b} = \frac{1}{2m\omega_2}\gamma_0(\not{\varepsilon}_1\not{\varepsilon}_2\not{k}_2)^\dagger\gamma_0 = \frac{1}{2m\omega_2}(\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1), \tag{4.61}$$

we can evaluate the traces $T_{\rm aa}$, $T_{\rm ab}$, $T_{\rm ba}$ and $T_{\rm bb}$. These have been given as Problem 4.2 and its solution given at the end of this chapter. Here, we only give the final results.

$$T_{\rm aa} = \frac{2}{m\omega_1}\left\{m\omega_2 + 2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2\right\}. \tag{4.62}$$

$$T_{\rm bb} = \frac{2}{m\omega_2}\left\{m\omega_1 - 2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2\right\}. \tag{4.63}$$

$$T_{\rm ab} = 2\left\{2(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2 - 1\right\} - \frac{2}{m\omega_1}(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2 + \frac{2}{m\omega_2}(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2. \tag{4.64}$$

$$T_{\rm ba} = T_{\rm ab}. \tag{4.65}$$

The term $T_{\rm bb}$ can be obtained from $T_{\rm aa}$ by using the crossing symmetry ($\mathbf{k}_1 \longrightarrow -\mathbf{k}_2$, $\boldsymbol{\varepsilon}_1 \longrightarrow \boldsymbol{\varepsilon}_2$, $\omega_1 \longrightarrow \omega_2$ but the term $T_{\rm ab}$ is invariant under the transformation representing crossing symmetry and it is equal to $T_{\rm ba}$.

Substituting (4.62) - (4.65) into Eq. (4.52), we obtain

$$\overline{\sum}|\bar{u}_2\mathcal{O}u_1|^2 = \frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} + 4(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2 - 2. \tag{4.66}$$

Consequently the matrix element square, given by Eq. (4.50) becomes

$$|\overline{\mathcal{M}}|^2 = 16\pi^2e^4\left\{\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} + 4(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2 - 2\right\}. \tag{4.67}$$

$$\text{Transition probability per unit time} = 2\pi(\Pi N)^{-1}|\overline{\mathcal{M}}|^2\,\rho_f, \tag{4.68}$$

where ΠN denotes the normalization factors that should be included for the incoming and outgoing particles and ρ_f, the density of final states. On dividing the transition probability per unit time (transition rate) by

incident flux, we obtain the cross section. Since the incident flux is c which is unity in natural units, we get

$$\begin{aligned} d\sigma &= \frac{2\pi}{2E_1 2E_2 2\omega_1 2\omega_2} |\overline{\mathcal{M}}|^2 \, \rho_f \\ &= \frac{2\pi}{2m 2E_2 2\omega_1 2\omega_2} |\overline{\mathcal{M}}|^2 \, \rho_f, \end{aligned} \tag{4.69}$$

since $E_1 = m$ in the laboratory system in which the initial electron is at rest.

4.2.4 Density of final states

For the two-particle final states, the density of states is given by (*vide* Appendix B)

$$\rho_f = (2\pi)^{-3} E_a E_b \frac{p_a^3 d\Omega_a}{E p_a^2 - E_a (\boldsymbol{p} \cdot \boldsymbol{p}_a)}, \tag{4.70}$$

where $E = E_a + E_b$ and $\boldsymbol{p} = \boldsymbol{p}_a + \boldsymbol{p}_b$. Identifying the particle a with the outgoing photon and particle b with the scattered electron, we get

$$\begin{aligned} E_a &= \omega_2; \qquad E_b = E_2; \qquad E = \omega_2 + E_2 = m + \omega_1. \\ &\qquad\qquad\qquad\qquad\quad \text{(since energy is conserved)} \\ p_a &= \omega_2; \qquad p_b = p_2; \qquad p = \omega_2 + p_2 = \omega_1. \\ &\qquad\qquad\qquad\qquad\quad \text{(since momentum is conserved)} \end{aligned}$$

Substituting these values in Eq. (4.70), we get

$$\begin{aligned} \rho_f &= (2\pi)^{-3} \omega_2 E_2 \frac{\omega_2^3 d\Omega_\omega}{(\omega_2 + E_2)\omega_2^2 - \omega_2(\omega_1 \omega_2 \cos\theta)} \\ &= (2\pi)^{-3} \frac{E_2 \omega_2^3 d\Omega_\omega}{(m + \omega_1)\omega_2 - \omega_1 \omega_2 \cos\theta} \\ &= (2\pi)^{-3} \frac{E_2 \omega_2^3 d\Omega_\omega}{m\omega_1}. \end{aligned} \tag{4.71}$$

The last step is obtained using the relation (4.35) for the Compton frequency shift.

4.2.5 Klein-Nishina formula

Substituting Eqs. (4.67) and (4.71) into Eq. (4.69), we obtain the Klein-Nishina formula for the Compton scattering cross section.

$$\frac{d\sigma}{d\Omega_\omega} = \frac{e^4}{4m^2} \left(\frac{\omega_2}{\omega_1}\right)^2 \left[\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} + 4(\boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2)^2 - 2\right]. \tag{4.72}$$

For simplicity, we have derived the formula for real polarization vectors which corresponds to linear polarization. If the polarization vectors are complex, then one has to replace $(\boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2)^2$ by $(\boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2^*)^2$ in Eq. (4.72). If the polarizations of the incident and final photons are not observed, then the appropriate cross section is obtained by averaging the initial polarization states and summing over the final polarization states. The result of this operation is to replace $(\boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_2)^2$ by $\frac{1}{2}(1 + \cos^2\theta)$, while the terms independent of $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are multiplied by 2.

$$\frac{d\sigma}{d\Omega_\omega} = \frac{e^4}{2m^2}\left(\frac{\omega_2}{\omega_1}\right)^2 \left[\frac{\omega_2}{\omega_1} + \frac{\omega_1}{\omega_2} - \sin^2\theta\right]. \tag{4.73}$$

Total cross section

Equation (4.73) gives the differential cross section. To obtain the total cross section, we need to perform the angular integration. A word of caution is necessary. The quantity ω_2 occurring in expression (4.73) depends on the scattering angle θ as given by Eq. (4.35). Multiplying Eq. (4.35) by ω_1/m, we obtain

$$\frac{\omega_1}{\omega_2} = 1 + \frac{\omega_1}{m}(1 - \cos\theta) = 1 + r(1 - \cos\theta), \tag{4.74}$$

$$\frac{\omega_2}{\omega_1} = \frac{1}{1 + \frac{\omega_1}{m}(1 - \cos\theta)} = \frac{1}{1 + r(1 - \cos\theta)}, \tag{4.75}$$

where $r = \frac{\omega_1}{m}$. The angular integration is performed taking into account the angular dependence of ω_2 (*vide* solved problem 4.3) to obtain the total cross section.

$$\sigma_{\text{total}} = \frac{2\pi e^4}{m^2}\left[\frac{1+r}{(1+2r)^2} + \frac{2}{r^2} - \frac{2(1+r) - r^2}{2r^3}\ln(1+2r)\right]. \tag{4.76}$$

Non-relativistic limit

In the non-relativistic limit of photon-electron scattering, when the energy of the incident photon is very small, $\omega_1 << m$, then $\omega_2 \approx \omega_1$ and the scattering can be treated as elastic. In this limiting case, Eq. (4.73) simplifies to

$$\frac{d\sigma}{d\Omega_\omega} = \frac{e^4}{2m^2}(2 - \sin^2\theta) = \frac{e^4}{2m^2}(1 + \cos^2\theta). \tag{4.77}$$

Performing the angular integration, we obtain the total cross section.

$$\sigma = \frac{e^4}{2m^2}\left(4\pi + \frac{4\pi}{3}\right) = \frac{8\pi e^4}{3m^2} = 0.665 \times 10^{-24}\ \text{cm}^2. \tag{4.78}$$

This is the Thomson scattering formula and can be considered as the non-relativistic limit of Compton scattering for unpolarized photons.

Extreme relativistic limit

In the other extreme relativistic limit when the incident photon energy is extremely large ($\omega_1 >> m$), then Eq. (4.35) reduces to

$$\omega_2 \approx \frac{m}{1-\cos\theta}$$

and the differential cross section (4.73) becomes

$$\frac{d\sigma}{d\Omega_\omega} = \frac{e^4}{2m^2}\frac{\omega_2}{\omega_1} = \frac{e^4}{2m\omega_1(1-\cos\theta)} = \frac{e^4}{4m\omega_1 \sin^2(\theta/2)}. \tag{4.79}$$

The total cross section in the extreme relativistic limit follows from Eq. (4.76).

$$\sigma = \frac{\pi e^4}{m\omega_1}\left\{\ln\frac{2\omega_1}{m} + \frac{1}{2}\right\}. \tag{4.80}$$

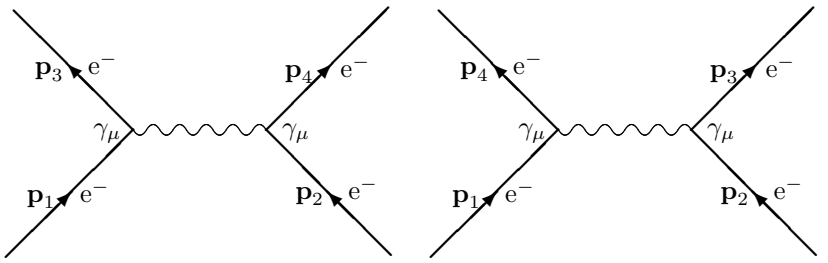
4.3 Electron-electron Scattering

Electron-electron scattering in the lowest order can be described by an exchange of one photon. Since electrons are identical particles, it is not possible to distinguish the two electrons in the final state. So, the scattering process is represented by two Feynman diagrams (a) and (b) as shown in Fig. 4.4. In Fig. 4.4(b), the four-momenta of the outgoing electrons are exchanged.

The matrix element for the process can easily be written down in the momentum representation by using the Feynman rules. For the sake of brevity, we denote the spinors $u(\mathbf{p}_1), u(\mathbf{p}_2)\cdots$ by $u_1, u_2, \cdots$.

$$\mathcal{M} = i4\pi e^2\left[\frac{(\bar{u}_4\gamma_\mu u_2)(\bar{u}_3\gamma_\mu u_1)}{(\mathbf{p}_1-\mathbf{p}_3)^2} - \frac{(\bar{u}_3\gamma_\mu u_2)(\bar{u}_4\gamma_\mu u_1)}{(\mathbf{p}_1-\mathbf{p}_4)^2}\right], \tag{4.81}$$

where summation over μ is implied. Thus, the longitudinal and transverse waves of the photon are taken into account in a relativistically invariant



(a) (b)

Figure 4.4: Electron-electron scattering. $\mathbf{p}_1$ and $\mathbf{p}_2$ denote the four-momenta of the electrons before collision and $\mathbf{p}_3$ and $\mathbf{p}_4$ after collision. Due to the indistinguishability of the two electrons in the final state, the exchange diagram (b) should also be considered.

way. The first term in Eq. (4.81) corresponds to the diagram (a) and the second term corresponds to the exchange diagram (b). The negative sign between the two terms arises from the antisymmetry of the two-electron wave function. To find the transition probability, we have to find the absolute square[9] of the matrix element (4.81). If the electron spins are not observed, then a sum over the final spin states and an average over the initial spin states have to be taken. This introduces a factor 1/4.

$$\begin{aligned}
|\overline{\mathcal{M}}|^2 &= \frac{1}{4}\mathcal{M}\mathcal{M}^\dagger \\
&= 4\pi^2 e^4 \sum_{\text{spins}} \left[\frac{(\bar{u}_4\gamma_\mu u_2)(\bar{u}_3\gamma_\mu u_1)}{(\mathbf{p}_1-\mathbf{p}_3)^2} - \frac{(\bar{u}_3\gamma_\mu u_2)(\bar{u}_4\gamma_\mu u_1)}{(\mathbf{p}_1-\mathbf{p}_4)^2} \right] \\
&\quad \times \left[\frac{(\bar{u}_4\gamma_\nu u_2)(\bar{u}_3\gamma_\nu u_1)}{(\mathbf{p}_1-\mathbf{p}_3)^2} - \frac{(\bar{u}_3\gamma_\nu u_2)(\bar{u}_4\gamma_\nu u_1)}{(\mathbf{p}_1-\mathbf{p}_4)^2} \right]^\dagger \\
&= 4\pi^2 e^4 \left[\frac{T_{\text{aa}}}{(\mathbf{p}_1-\mathbf{p}_3)^4} + \frac{T_{\text{bb}}}{(\mathbf{p}_1-\mathbf{p}_4)^4} - \frac{T_{\text{ab}}+T_{\text{ba}}}{(\mathbf{p}_1-\mathbf{p}_3)^2(\mathbf{p}_1-\mathbf{p}_4)^2} \right], \qquad (4.82)
\end{aligned}$$

[9] In finding the absolute square of the matrix element, the summation index μ has to be changed to ν in the hermitian conjugate of the matrix element, since

$$\sum_\mu a_\mu \Big)^2 = \sum_\mu a_\mu \Big) \sum_\nu a_\nu^* \Big) = \sum_{\mu\nu} a_\mu a_\nu^*.$$

where

$$
\begin{aligned}
T_{\rm aa} &= \sum_{\rm spins}\left[(\bar{u}_3\gamma_\mu u_1)(\bar{u}_3\gamma_\nu u_1)^\dagger(\bar{u}_4\gamma_\mu u_2)(\bar{u}_4\gamma_\nu u_2)^\dagger\right] \\
&= {\rm Tr}\left[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_3+m)\right]{\rm Tr}\left[\gamma_\mu(\not{p}_2+m)\gamma_\nu(\not{p}_4+m)\right], \quad (4.83)\\
T_{\rm bb} &= \sum_{\rm spins}\left[(\bar{u}_4\gamma_\mu u_1)(\bar{u}_4\gamma_\nu u_1)^\dagger(\bar{u}_3\gamma_\mu u_2)(\bar{u}_3\gamma_\nu u_2)^\dagger\right] \\
&= {\rm Tr}\left[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_4+m)\right]{\rm Tr}\left[\gamma_\mu(\not{p}_2+m)\gamma_\nu(\not{p}_3+m)\right], \quad (4.84)\\
T_{\rm ab} &= \sum_{\rm spins}\left[(\bar{u}_3\gamma_\mu u_1)(\bar{u}_4\gamma_\nu u_1)^\dagger(\bar{u}_4\gamma_\mu u_2)(\bar{u}_3\gamma_\nu u_2)^\dagger\right] \\
&= {\rm Tr}\left[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_4+m)\gamma_\mu(\not{p}_2+m)\gamma_\nu(\not{p}_3+m)\right], \quad (4.85)\\
T_{\rm ba} &= \sum_{\rm spins}\left[(\bar{u}_4\gamma_\mu u_1)(\bar{u}_3\gamma_\nu u_1)^\dagger(\bar{u}_3\gamma_\mu u_2)(\bar{u}_4\gamma_\nu u_2)^\dagger\right] \\
&= {\rm Tr}\left[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_3+m)\gamma_\mu(\not{p}_2+m)\gamma_\nu(\not{p}_4+m)\right]. \quad (4.86)
\end{aligned}
$$

4.3.1 Evaluation of traces

To evaluate the traces $T_{\rm aa}$ and $T_{\rm bb}$, we use the following trace formula:

$$
{\rm Tr}[\gamma_\mu(\not{a}+m)\gamma_\nu(\not{b}+m)] = 4\left[a_\mu b_\nu + b_\mu a_\nu - g_{\mu\nu}(\mathbf{a}\cdot\mathbf{b}-m^2)\right]. \quad (4.87)
$$

Using Eq. (4.87), we find after simplification

$$
\begin{aligned}
T_{\rm aa} &= 16\left[p_{1\mu}p_{3\nu}+p_{3\mu}p_{1\nu}-g_{\mu\nu}(\mathbf{p}_1\cdot\mathbf{p}_3-m^2)\right] \\
&\quad\times\left[p_{2\mu}p_{4\nu}+p_{4\mu}p_{2\nu}-g_{\mu\nu}(\mathbf{p}_2\cdot\mathbf{p}_4-m^2)\right] \\
&= 32\left[(\mathbf{p}_1\cdot\mathbf{p}_2)^2+(\mathbf{p}_1\cdot\mathbf{p}_4)^2+2m^2(m^2-\mathbf{p}_1\cdot\mathbf{p}_3)\right], \quad (4.88)
\end{aligned}
$$

where we have used the relations $g_{\mu\nu}g_{\mu\nu}=4$ and

$$
\mathbf{p}_1\cdot\mathbf{p}_2=\mathbf{p}_3\cdot\mathbf{p}_4; \qquad \mathbf{p}_1\cdot\mathbf{p}_3=\mathbf{p}_2\cdot\mathbf{p}_4; \qquad \mathbf{p}_1\cdot\mathbf{p}_4=\mathbf{p}_2\cdot\mathbf{p}_3;
$$

which are obtained from the energy-momentum conservation law.

$$
\mathbf{p}_1+\mathbf{p}_2=\mathbf{p}_3+\mathbf{p}_4.
$$

Following the same method, $T_{\rm bb}$ can be evaluated. But, we can write down the result by an inspection of Eq. (4.88) by interchanging $\mathbf{p}_3 \leftrightarrow \mathbf{p}_4$.

$$
T_{\rm bb} = 32\left[(\mathbf{p}_1\cdot\mathbf{p}_2)^2+(\mathbf{p}_1\cdot\mathbf{p}_3)^2+2m^2(m^2-\mathbf{p}_1\cdot\mathbf{p}_4)\right], \quad (4.89)
$$

The evaluation of $T_{\rm ab}$ is a little involved. Since the trace of a product of odd number of γ matrices vanishes, we need to take only the terms involving even number of γ matrices.

$$\begin{aligned} T_{\rm ab} &= \mathrm{Tr}\,[\gamma_\mu \not{p}_1 \gamma_\nu \not{p}_4 \gamma_\mu \not{p}_2 \gamma_\nu \not{p}_3] \\ &\quad + m^2\,\mathrm{Tr}\,[\gamma_\mu \not{p}_1 \gamma_\nu \not{p}_4 \gamma_\mu \gamma_\nu + \gamma_\mu \not{p}_1 \gamma_\nu \gamma_\mu \not{p}_2 \gamma_\nu + \gamma_\mu \not{p}_1 \gamma_\nu \gamma_\mu \gamma_\nu \not{p}_3 \\ &\qquad + \gamma_\mu \gamma_\nu \not{p}_4 \gamma_\mu \not{p}_2 \gamma_\nu + \gamma_\mu \gamma_\nu \not{p}_4 \gamma_\mu \gamma_\nu \not{p}_3 + \gamma_\mu \gamma_\nu \gamma_\mu \not{p}_2 \gamma_\nu \not{p}_3] \\ &\quad + m^4\,\mathrm{Tr}\,[\gamma_\mu \gamma_\nu \gamma_\mu \gamma_\nu]\,. \end{aligned} \tag{4.90}$$

Using the anticommutation relations of γ matrices, we can prove the following formulae involving a product of γ matrices of repeated index[10] (*vide* Problem 4.1):

$$\gamma_\mu \gamma_\mu = 4; \qquad \gamma_\mu \gamma_\nu \gamma_\mu = -2\gamma_\nu; \tag{4.91}$$

$$\gamma_\mu \gamma_\nu \gamma_\rho \gamma_\mu = 4 g_{\nu\rho}; \qquad \gamma_\mu \gamma_\nu \gamma_\rho \gamma_\lambda \gamma_\mu = -2\gamma_\lambda \gamma_\rho \gamma_\nu. \tag{4.92}$$

Using the above formulae, we get[11]

$$\begin{aligned} \underbrace{\gamma_\mu \not{p}_1 \gamma_\nu \not{p}_4 \gamma_\mu} \not{p}_2 \gamma_\nu \not{p}_3 &= -2 \not{p}_4 \underbrace{\gamma_\nu \not{p}_1 \not{p}_2 \gamma_\nu} \not{p}_3 \\ &= -8(\mathbf{p}_1 \cdot \mathbf{p}_2) \not{p}_4 \not{p}_3; \end{aligned} \tag{4.93}$$

$$\begin{aligned} \mathrm{Tr}(\gamma_\mu \not{p}_1 \gamma_\nu \not{p}_4 \gamma_\mu \not{p}_2 \gamma_\nu \not{p}_3) &= -8(\mathbf{p}_1 \cdot \mathbf{p}_2)\mathrm{Tr}(\not{p}_4 \not{p}_3) \\ &= -32(\mathbf{p}_1 \cdot \mathbf{p}_2)(\mathbf{p}_3 \cdot \mathbf{p}_4); \end{aligned} \tag{4.94}$$

$$\underbrace{\gamma_\mu \not{p}_1 \gamma_\nu \not{p}_4 \gamma_\mu} \gamma_\nu = -2 \not{p}_4 \underbrace{\gamma_\nu \not{p}_1 \gamma_\nu} = 4 \not{p}_4 \not{p}_1; \tag{4.95}$$

$$\mathrm{Tr}(\gamma_\mu \not{p}_1 \gamma_\nu \not{p}_4 \gamma_\mu \gamma_\nu) = 4\mathrm{Tr}(\not{p}_4 \not{p}_1) = 16(\mathbf{p}_1 \cdot \mathbf{p}_4); \tag{4.96}$$

$$\underbrace{\gamma_\mu \gamma_\nu \gamma_\mu} \gamma_\nu = -2\gamma_\nu \gamma_\nu, \qquad \text{since } \gamma_\mu \gamma_\nu \gamma_\mu = -2\gamma_\nu \tag{4.97}$$

$$\mathrm{Tr}\,(\gamma_\mu \gamma_\nu \gamma_\mu \gamma_\nu) = -2\mathrm{Tr}\,(\gamma_\mu)^2 = -8\mathrm{Tr}\,I \,=\, -32. \tag{4.98}$$

Substituting the above results in Eq. (4.90), we get

$$\begin{aligned} T_{\rm ab} &= -32(\mathbf{p}_1 \cdot \mathbf{p}_2)(\mathbf{p}_3 \cdot \mathbf{p}_4) + 16m^2\{\mathbf{p}_1 \cdot \mathbf{p}_4 + \mathbf{p}_1 \cdot \mathbf{p}_2 + \mathbf{p}_1 \cdot \mathbf{p}_3 \\ &\qquad + \mathbf{p}_2 \cdot \mathbf{p}_4 + \mathbf{p}_3 \cdot \mathbf{p}_4 + \mathbf{p}_2 \cdot \mathbf{p}_3\} - 32m^4 \\ &= -32(\mathbf{p}_1 \cdot \mathbf{p}_2)^2 + 32m^2\{\mathbf{p}_1 \cdot \mathbf{p}_4 + \mathbf{p}_1 \cdot \mathbf{p}_2 + \mathbf{p}_1 \cdot \mathbf{p}_3\} - 32m^4. \end{aligned} \tag{4.99}$$

[10] A summation over the repeated index is implied.

[11] Underbraces are used to highlight the factors that are considered in the step-by-step calculation for the benefit of the reader.

Equation (4.99) remains the same under the exchange $\mathbf{p}_3 \leftrightarrow \mathbf{p}_4$. So, the trace $T_{\rm ba} = T_{\rm ab}$. Hence

$$\begin{aligned} T_{\rm ab} + T_{\rm ba} &= 64\left[-(\mathbf{p}_1\cdot\mathbf{p}_2)^2 + m^2\{\mathbf{p}_1\cdot\mathbf{p}_4 + \mathbf{p}_1\cdot\mathbf{p}_2 + \mathbf{p}_1\cdot\mathbf{p}_3\} - m^4\right] \\ &= 64\left[-(\mathbf{p}_1\cdot\mathbf{p}_2)^2 + 2m^2(\mathbf{p}_1\cdot\mathbf{p}_2)\right]. \end{aligned} \tag{4.100}$$

The last step is obtained by invoking the energy-momentum conservation.

$$\begin{aligned} \mathbf{p}_1\cdot\mathbf{p}_4 &= \mathbf{p}_1\cdot(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{p}_3) \\ &= m^2 + \mathbf{p}_1\cdot\mathbf{p}_2 - \mathbf{p}_1\cdot\mathbf{p}_3. \end{aligned} \tag{4.101}$$

The denominators in Eq. (4.82) can be expressed in terms of scalar products of four vectors.

$$(\mathbf{p}_1 - \mathbf{p}_3)^2 = \mathbf{p}_1^2 + \mathbf{p}_3^2 - 2\mathbf{p}_1\cdot\mathbf{p}_3 = 2(m^2 - \mathbf{p}_1\cdot\mathbf{p}_3). \tag{4.102}$$

$$(\mathbf{p}_1 - \mathbf{p}_4)^2 = \mathbf{p}_1^2 + \mathbf{p}_4^2 - 2\mathbf{p}_1\cdot\mathbf{p}_4 = 2(m^2 - \mathbf{p}_1\cdot\mathbf{p}_4). \tag{4.103}$$

$$\begin{aligned} |\overline{\mathcal{M}}|^2 &= 32\pi^2 e^4 \left[\frac{\left\{(\mathbf{p}_1\cdot\mathbf{p}_2)^2 + (\mathbf{p}_1\cdot\mathbf{p}_4)^2 + 2m^2(m^2 - \mathbf{p}_1\cdot\mathbf{p}_3)\right\}}{(m^2 - \mathbf{p}_1\cdot\mathbf{p}_3)^2}\right. \\ &\quad + \frac{\left\{(\mathbf{p}_1\cdot\mathbf{p}_2)^2 + (\mathbf{p}_1\cdot\mathbf{p}_3)^2 + 2m^2(m^2 - \mathbf{p}_1\cdot\mathbf{p}_4)\right\}}{(m^2 - \mathbf{p}_1\cdot\mathbf{p}_4)^2} \\ &\quad \left. - \frac{2\left\{-(\mathbf{p}_1\cdot\mathbf{p}_2)^2 + 2m^2(\mathbf{p}_1\cdot\mathbf{p}_2)\right\}}{(m^2 - \mathbf{p}_1\cdot\mathbf{p}_3)(m^2 - \mathbf{p}_1\cdot\mathbf{p}_4)}\right]. \end{aligned} \tag{4.104}$$

Equation (4.104) is valid in all frames of reference, both in laboratory and centre of momentum frames.

4.3.2 In the centre of momentum frame

Let us choose the centre of momentum (c.m.) frame for evaluation of the cross section. It describes the elastic scattering of two electrons as shown in Fig. 4.5 with the three-momenta of incident electrons equal and opposite ($\boldsymbol{p}_2 = -\boldsymbol{p}_1$). Similarly, the scattered electrons have also their three-momenta equal and opposite ($\boldsymbol{p}_4 = -\boldsymbol{p}_3$). Also the energies of the incident and scattered electrons are equal.

$$|\boldsymbol{p}_1| = |\boldsymbol{p}_2| = |\boldsymbol{p}_3| = |\boldsymbol{p}_4| = p \quad \text{and} \quad E_1 = E_2 = E_3 = E_4 = E. \tag{4.105}$$

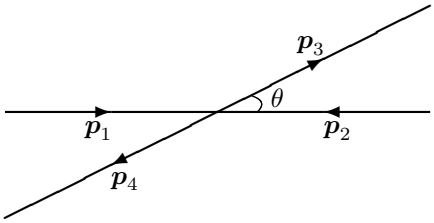


Figure 4.5: Electron-electron scattering in c.m. frame. $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ denote the three-momenta before collision and $\boldsymbol{p}_3$ and $\boldsymbol{p}_4$ denote the three-momenta after collision, θ being the angle of scattering.

Now we can write the scalar product of the four-vectors and the other quantities in terms of E, p and the angle of scattering θ.

$$\begin{aligned}
\mathbf{p}_1 \cdot \mathbf{p}_2 &= E_1 E_2 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_2 = E^2 + p^2. && (4.106)\\
\mathbf{p}_1 \cdot \mathbf{p}_3 &= E_1 E_3 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_3 = E^2 - p^2 \cos\theta. && (4.107)\\
\mathbf{p}_1 \cdot \mathbf{p}_4 &= E_1 E_4 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_4 = E^2 + p^2 \cos\theta. && (4.108)\\
(\mathbf{p}_1 - \mathbf{p}_3)^2 &= 2m^2(1 - \mathbf{p}_1 \cdot \mathbf{p}_3) = -2p^2(1 - \cos\theta). && (4.109)\\
(\mathbf{p}_1 - \mathbf{p}_4)^2 &= 2m^2(1 - \mathbf{p}_1 \cdot \mathbf{p}_4) = -2p^2(1 + \cos\theta). && (4.110)
\end{aligned}$$

Substituting the results (4.106) - (4.110) into Eqs. (4.104), we obtain an expression for $|\overline{\mathcal{M}}|^2$ defined in Eq. (4.82).

$$\begin{aligned}
|\overline{\mathcal{M}}|^2 = \frac{32\pi^2 e^4}{p^4} \Bigg[& \frac{\left\{(E^2+p^2)^2 + (E^2 + p^2\cos\theta)^2 - 2m^2p^2(1-\cos\theta)\right\}}{(1-\cos\theta)^2} \\
& + \frac{\left\{(E^2+p^2)^2 + (E^2 - p^2\cos\theta)^2 - 2m^2p^2(1+\cos\theta)\right\}}{(1+\cos\theta)^2} \\
& + \frac{2\left\{(E^2+p^2)^2 - 2m^2(E^2+p^2)\right\}}{\sin^2\theta} \Bigg]. \qquad (4.111)
\end{aligned}$$

Differential cross section in c.m. frame

The cross section is given by

$$d\sigma = \frac{2\pi}{v}(\Pi N)^{-1} |\overline{\mathcal{M}}|^2 \rho_f, \qquad (4.112)$$

where $(\Pi N)^{-1}$ denotes the normalization factor

$$(\Pi N)^{-1} = \frac{1}{2E_1 2E_2 2E_3 2E_4} = \left(\frac{1}{2E}\right)^4$$

and ρ_f is the density of final states for the two-particle system (*vide* Appendix).

$$\begin{aligned}\rho_f &= \frac{1}{(2\pi)^3} E_3 E_4 \frac{p_3^3 d\Omega}{(E_3+E_4)p_3^2 - E_3(\boldsymbol{P}\cdot\boldsymbol{p}_3)} \\ &= \frac{1}{(2\pi)^3} E^2 \frac{p^3 d\Omega}{2Ep^2} = \frac{Ep}{2(2\pi)^3} d\Omega, \end{aligned} \tag{4.113}$$

where $\boldsymbol{P} = \boldsymbol{p}_3 + \boldsymbol{p}_4$ denotes the total momentum of the two-particle system which is zero in the c.m. frame. The incident flux depends upon the relative velocity v of the two colliding particles and in the c.m. system, it is equal to twice the velocity of either of the colliding partners which are of equal mass.

$$v = 2\frac{p}{E}.$$

Collecting all these results, we arrive at an expression for the differential cross section.

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \frac{2\pi}{2(p/E)} \left(\frac{1}{2E}\right)^4 |\overline{\mathcal{M}}|^2 \frac{Ep}{2(2\pi)^3} \\ &= \frac{1}{64E^2} \frac{1}{(2\pi)^2} |\overline{\mathcal{M}}|^2. \end{aligned} \tag{4.114}$$

Equation (4.114) together with Eq. (4.111) gives the electron-electron scattering cross section in c.m. frame. The electron-electron scattering is sometimes known as Møller scattering since it was Møller[12] who first deduced the electron-electron scattering cross section in the laboratory frame using relativistic quantum mechanics.

4.4 Electron-positron Scattering

Electron-positron scattering is often referred to as Bhabha scattering since it was he who first deduced the relativistic cross section for this process[13]. Electron-positron scattering is similar electron-electron scattering but the arrows representing the positron are reversed, indicating that the positron travels backward in time according to Feynman's positron theory. In this case also, there are two Feynman diagrams, one indicating an exchange of a photon between electron and positron and the other indicating electron-positron pair annihilation into a photon which subsequently creates an

[12] C. Møller, *Ann. der. Phys.*, **85**, 711 (1932).

[13] H. J. Bhabha, *Proc. Roy. Soc (London)***A154**, 195 (1936).

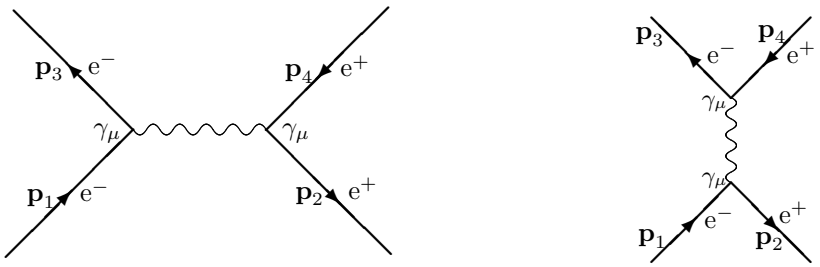


Figure 4.6: Electron-positron scattering. Diagram (a) represents the exchange of one photon between electron and positron and Diagram (b) denotes the $e^+\,e^-$ pair annihilation and subsequent creation. Both the diagrams contribute to the process.

electron-positron pair. Both the diagrams contribute to the Feynman amplitude for the process. Using the short-hand notation $u_1, \bar{v}_2, \bar{u}_3, v_4$ for the spinors $u(\mathbf{p}_1), \bar{v}(\mathbf{p}_2)$, $\bar{u}(\mathbf{p}_3), v(\mathbf{p}_4)$, we write down the matrix element for the electron-positron scattering.

$$\mathcal{M} = i4\pi e^2 \left[\frac{(\bar{u}_3\gamma_\mu u_1)(\bar{v}_2\gamma_\mu v_4)}{(\mathbf{p}_1 - \mathbf{p}_3)^2} - \frac{(\bar{v}_2\gamma_\mu u_1)(\bar{u}_3\gamma_\mu v_4)}{(\mathbf{p}_1 + \mathbf{p}_2)^2} \right], \tag{4.115}$$

The relative negative sign between the two amplitudes arises from fermion exchange - the initial e^+ (negative energy electron) is exchanged with the final e^-. To calculate the cross section for unpolarized particles, we need to find the absolute square of the above amplitude (4.115) by summing over the final spin states and averaging over the initial spin states.

$$\begin{aligned}
|\overline{\mathcal{M}}|^2 &= \frac{1}{4}\mathcal{M}\mathcal{M}^\dagger \\
&= 4\pi^2 e^4 \sum_{\text{spins}} \left[\frac{\{(\bar{u}_3\gamma_\mu u_1)(\bar{u}_1\gamma_\nu u_3)\}\,\{(\bar{v}_2\gamma_\mu v_4)(\bar{v}_4\gamma_\nu v_2)\}}{(\mathbf{p}_1 - \mathbf{p}_3)^4} \right. \\
&\qquad + \frac{\{(\bar{v}_2\gamma_\mu u_1)(\bar{u}_1\gamma_\nu v_2)\}\,\{(\bar{u}_3\gamma_\mu v_4)(\bar{v}_4\gamma_\nu u_3)\}}{(\mathbf{p}_1 + \mathbf{p}_2)^4} \\
&\qquad - \frac{\{(\bar{u}_3\gamma_\mu u_1)(\bar{u}_1\gamma_\nu v_2)(\bar{v}_2\gamma_\mu v_4)(\bar{v}_4\gamma_\nu u_3)\}}{(\mathbf{p}_1 - \mathbf{p}_3)^2(\mathbf{p}_1 + \mathbf{p}_2)^2} \\
&\qquad \left. - \frac{\{(\bar{v}_2\gamma_\mu u_1)(\bar{u}_1\gamma_\nu u_3)(\bar{u}_3\gamma_\mu v_4)(\bar{v}_4\gamma_\nu v_2)\}}{(\mathbf{p}_1 - \mathbf{p}_3)^2(\mathbf{p}_1 + \mathbf{p}_2)^2} \right] \\
&= 4\pi^2 e^4 \left[\frac{T_{\text{aa}}}{(\mathbf{p}_1 - \mathbf{p}_3)^4} + \frac{T_{\text{bb}}}{(\mathbf{p}_1 + \mathbf{p}_2)^4} - \frac{T_{\text{ab}} + T_{\text{ba}}}{(\mathbf{p}_1 - \mathbf{p}_3)^2(\mathbf{p}_1 + \mathbf{p}_2)^2} \right],
\end{aligned} \tag{4.116}$$

where

$$
\begin{aligned}
T_{\rm aa} &= \sum_{\rm spins} [\{(\bar{u}_3\gamma_\mu u_1)(\bar{u}_1\gamma_\nu u_3)\}\{(\bar{v}_2\gamma_\mu v_4)(\bar{v}_4\gamma_\nu v_2)\}] \\
&= {\rm Tr}\,[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_3+m)]\,{\rm Tr}\,[\gamma_\mu(\not{p}_4-m)\gamma_\nu(\not{p}_2-m)]\,; \quad (4.117) \\
T_{\rm bb} &= \sum_{\rm spins} [\{(\bar{v}_2\gamma_\mu u_1)(\bar{u}_1\gamma_\nu v_2)\}\{(\bar{u}_3\gamma_\mu v_4)(\bar{v}_4\gamma_\nu u_3)\}] \\
&= {\rm Tr}\,[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_2-m)]\,{\rm Tr}\,[\gamma_\mu(\not{p}_4-m)\gamma_\nu(\not{p}_3+m)]\,; \quad (4.118) \\
T_{\rm ab} &= \sum_{\rm spins} [(\bar{u}_3\gamma_\mu u_1)(\bar{u}_1\gamma_\nu v_2)(\bar{v}_2\gamma_\mu v_4)(\bar{v}_4\gamma_\nu u_3)] \\
&= {\rm Tr}\,[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_2-m)\gamma_\mu(\not{p}_4-m)\gamma_\nu(\not{p}_3+m)]\,; \quad (4.119) \\
T_{\rm ba} &= \sum_{\rm spins} [(\bar{v}_2\gamma_\mu u_1)(\bar{u}_1\gamma_\nu u_3)(\bar{u}_3\gamma_\mu v_4)(\bar{v}_4\gamma_\nu v_2)] \\
&= {\rm Tr}\,[\gamma_\mu(\not{p}_1+m)\gamma_\nu(\not{p}_3+m)\gamma_\mu(\not{p}_4-m)\gamma_\nu(\not{p}_2-m)]\,. \quad (4.120)
\end{aligned}
$$

These traces can be calculated by following the same procedure used in the previous section on electron-electron scattering. By an inspection, it is found that the traces can be written down directly from Eqs. (4.88), (4.89), (4.99) and (4.100) by interchanging $\mathbf{p}_2 \leftrightarrow -\mathbf{p}_4$.

$$
\begin{aligned}
T_{\rm aa} &= 32\left[(\mathbf{p}_1\cdot\mathbf{p}_2)^2+(\mathbf{p}_1\cdot\mathbf{p}_4)^2+2m^2(m^2-\mathbf{p}_1\cdot\mathbf{p}_3)\right], \quad (4.121) \\
T_{\rm bb} &= 32\left[(\mathbf{p}_1\cdot\mathbf{p}_4)^2+(\mathbf{p}_1\cdot\mathbf{p}_3)^2+2m^2(m^2+\mathbf{p}_1\cdot\mathbf{p}_2)\right], \quad (4.122) \\
T_{\rm ab} &= -32(\mathbf{p}_1\cdot\mathbf{p}_4)^2-32m^2\{\mathbf{p}_1\cdot\mathbf{p}_2+\mathbf{p}_1\cdot\mathbf{p}_4-\mathbf{p}_1\cdot\mathbf{p}_3\}-32m^4. \quad (4.123) \\
T_{\rm ba} &= -32(\mathbf{p}_1\cdot\mathbf{p}_4)^2-32m^2\{\mathbf{p}_1\cdot\mathbf{p}_2+\mathbf{p}_1\cdot\mathbf{p}_4-\mathbf{p}_1\cdot\mathbf{p}_3\}-32m^4. \quad (4.124)
\end{aligned}
$$

Differential cross section in c.m. frame

Till now, we have not specified the frames of reference and so the expressions (4.121) - (4.124) are valid in all frames of reference. Now, we shall express the scalar products of four-vectors in c.m. frame as given by Eqs. (4.106) - (4.110) and substitute them in Eq. (4.116) to get $|\overline{\mathcal{M}}|^2$.

$$
\begin{aligned}
|\overline{\mathcal{M}}|^2 = 32\pi^2e^4\Bigg[&\frac{(E^2+p^2)^2+(E^2+p^2\cos\theta)^2-2m^2p^2(1-\cos\theta)}{p^4(1-\cos\theta)^2} \\
&+\frac{E^4+p^4\cos^2\theta+2m^2E^2}{2E^4} \\
&-\frac{(E^2+p^2\cos\theta)^2+2m^2(E^2+p^2\cos\theta)}{E^2p^2(1-\cos\theta)}\Bigg]. \quad (4.125)
\end{aligned}
$$

By substituting expression (4.125) in Eq. (4.114), we obtain the differential cross section for electron - positron scattering.

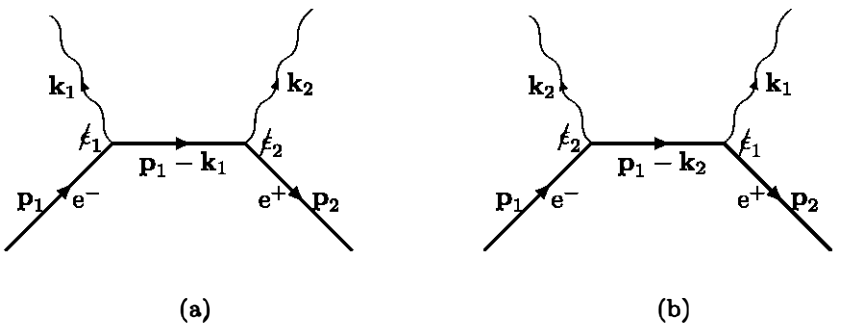
$$\frac{d\sigma}{d\Omega} = \frac{e^4}{8E^2}\left[\frac{(E^2+p^2)^2+(E^2+p^2\cos\theta)^2-2m^2p^2(1-\cos\theta)}{p^4(1-\cos\theta)^2} + \frac{E^4+p^4\cos^2\theta+2m^2E^2}{2E^4} - \frac{(E^2+p^2\cos\theta)^2+2m^2(E^2+p^2\cos\theta)}{E^2p^2(1-\cos\theta)}\right], \quad (4.126)$$

where e^2 is the fine structure constant α. In the high energy limit ($E >> m$), Eq. (4.126) reduces to

$$\frac{d\sigma}{d\Omega} = \frac{e^4}{8E^2}\left[\frac{1+\cos^4(\theta/2)}{\sin^4(\theta/2)} + \frac{1+\cos^2\theta}{2} - \frac{2\cos^4(\theta/2)}{\sin^2(\theta/2)}\right]. \quad (4.127)$$

4.5 Electron-positron Pair Annihilation into two Photons

We have observed earlier that the first order process of electron-positron pair annihilation into a single photon is unphysical and so we need to consider necessarily a second order process of pair annihilation into two photons. This is depicted by Feynman diagrams in Fig. 4.7. The diagram (b) represents the exchange diagram. The Feynman amplitudes for



(a) (b)

Figure 4.7: Electron-positron pair annihilation into two photons. The internal line connecting the two vertices denotes a virtual particle.

diagrams (a) and (b) can be written as

$$\mathcal{M}_{\rm a} = -i4\pi e^2 \left(\bar{v}_2 \not{\epsilon}_2 \frac{1}{\not{p}_1 - \not{k}_1 - m} \not{\epsilon}_1 u_1 \right), \tag{4.128}$$

$$\mathcal{M}_{\rm b} = -i4\pi e^2 \left(\bar{v}_2 \not{\epsilon}_1 \frac{1}{\not{p}_1 - \not{k}_2 - m} \not{\epsilon}_2 u_1 \right), \tag{4.129}$$

$$\mathcal{M} = \mathcal{M}_{\rm a} + \mathcal{M}_{\rm b} = -i4\pi e^2 \left\{ \bar{v}_2 (\mathcal{O}_{\rm a} + \mathcal{O}_{\rm b}) u_1 \right\}, \tag{4.130}$$

where

$$\mathcal{O}_{\rm a} = \not{\epsilon}_2 \frac{1}{\not{p}_1 - \not{k}_1 - m} \not{\epsilon}_1 = \not{\epsilon}_2 \frac{\not{p}_1 - \not{k}_1 + m}{(\not{p}_1 - \not{k}_1)^2 - m^2} \not{\epsilon}_1; \tag{4.131}$$

$$\mathcal{O}_{\rm b} = \not{\epsilon}_1 \frac{1}{\not{p}_1 - \not{k}_2 - m} \not{\epsilon}_2 = \not{\epsilon}_1 \frac{\not{p}_1 - \not{k}_2 + m}{(\not{p}_1 - \not{k}_2)^2 - m^2} \not{\epsilon}_2. \tag{4.132}$$

Let us now choose a coordinate system in which the electron is at rest. By this choice, the matrix element calculation is greatly simplified. The denominators in Eqs. (4.131) and (4.132) simplify to

$$\begin{aligned} (\not{p}_1 - \not{k}_1)^2 - m^2 &= \not{p}_1^2 + \not{k}_1^2 - 2\not{p}_1 \not{k}_1 - m^2 \\ &= \mathbf{p}_1^2 + \mathbf{k}_1^2 - 2\mathbf{p}_1 \cdot \mathbf{k}_1 - m^2 = -2m\omega_1; \end{aligned} \tag{4.133}$$

$$\begin{aligned} (\not{p}_1 - \not{k}_2)^2 - m^2 &= \not{p}_1^2 + \not{k}_2^2 - 2\not{p}_1 \not{k}_2 - m^2 \\ &= \mathbf{p}_1^2 + \mathbf{k}_2^2 - 2\mathbf{p}_1 \cdot \mathbf{k}_2 - m^2 = -2m\omega_2. \end{aligned} \tag{4.134}$$

The numerators can also be similarly simplified. The first term in the numerators of Eqs. (4.131) and (4.132) taken between the electron and positron spinors

$$\begin{aligned} \bar{v}_2 \not{\epsilon}_2 \underbrace{\not{p}_1 \not{\epsilon}_1} u_1 &= \bar{v}_2 \not{\epsilon}_2 (-\not{\epsilon}_1 \not{p}_1 + 2\mathbf{p}_1 \cdot \boldsymbol{\varepsilon}_1) u_1 \\ &= -\bar{v}_2 \not{\epsilon}_2 \not{\epsilon}_1 \not{p}_1 u_1 = -m \bar{v}_2 \not{\epsilon}_2 \not{\epsilon}_1 u_1, \end{aligned} \tag{4.135}$$

$$\begin{aligned} \bar{v}_2 \not{\epsilon}_1 \not{p}_1 \not{\epsilon}_2 u_1 &= \bar{v}_2 \not{\epsilon}_1 (-\not{\epsilon}_2 \not{p}_1 + 2\mathbf{p}_1 \cdot \boldsymbol{\varepsilon}_2) u_1 \\ &= -\bar{v}_2 \not{\epsilon}_1 \not{\epsilon}_2 \not{p}_1 u_1 = -m \bar{v}_2 \not{\epsilon}_1 \not{\epsilon}_2 u_1, \end{aligned} \tag{4.136}$$

cancels with the third term. Thus we obtain

$$\bar{v}_2 \mathcal{O}_{\rm a} u_1 = -\frac{\bar{v}_2 \not{\epsilon}_2 \not{k}_1 \not{\epsilon}_1 u_1}{2m\omega_1} = \frac{\bar{v}_2 \not{\epsilon}_2 \not{\epsilon}_1 \not{k}_1 u_1}{2m\omega_1}; \tag{4.137}$$

$$\bar{v}_2 \mathcal{O}_{\rm b} u_1 = -\frac{\bar{v}_2 \not{\epsilon}_1 \not{k}_2 \not{\epsilon}_2 u_1}{2m\omega_2} = \frac{\bar{v}_2 \not{\epsilon}_1 \not{\epsilon}_2 \not{k}_2 u_1}{2m\omega_2}. \tag{4.138}$$

Transition probability

If we are not interested in the spin states of the electron and positron, then an average has to be taken over the spin states in calculating the transition probability.

$$|\overline{\mathcal{M}}|^2 = 16\pi^2 e^4 \bar{\sum}|\bar{v}_2 \mathcal{O} u_1|^2, \tag{4.139}$$

where

$$\begin{aligned} \mathcal{O} &= \mathcal{O}_{\rm a} + \mathcal{O}_{\rm b} \\ &= \frac{\not\epsilon_2 \not\epsilon_1 \not k_1}{2m\omega_1} + \frac{\not\epsilon_1 \not\epsilon_2 \not k_2}{2m\omega_1}. \end{aligned} \tag{4.140}$$

Using the trace techniques, we get

$$\begin{aligned} \overline{\sum}|\bar{v}_2 \mathcal{O} u_1|^2 &= \frac{1}{4}{\rm Tr}\left[\mathcal{O}(\not p_1 + m)\tilde{\mathcal{O}}(\not p_2 - m)\right] \\ &= \frac{1}{4}{\rm Tr}\left[(\mathcal{O}_{\rm a} + \mathcal{O}_{\rm b})(\not p_1 + m)(\tilde{\mathcal{O}}_{\rm a} + \tilde{\mathcal{O}}_{\rm b})(\not p_2 - m)\right] \\ &= \frac{1}{4}\left[T_{\rm aa} + T_{\rm bb} + T_{\rm ab} + T_{\rm ba}\right], \end{aligned} \tag{4.141}$$

where

$$T_{\rm aa} = {\rm Tr}\left[\mathcal{O}_{\rm a}(\not p_1 + m)\tilde{\mathcal{O}}_{\rm a}(\not p_2 - m)\right], \tag{4.142}$$

$$T_{\rm bb} = {\rm Tr}\left[\mathcal{O}_{\rm b}(\not p_1 + m)\tilde{\mathcal{O}}_{\rm b}(\not p_2 - m)\right], \tag{4.143}$$

$$T_{\rm ab} = {\rm Tr}\left[\mathcal{O}_{\rm a}(\not p_1 + m)\tilde{\mathcal{O}}_{\rm b}(\not p_2 - m)\right], \tag{4.144}$$

$$T_{\rm ba} = {\rm Tr}\left[\mathcal{O}_{\rm b}(\not p_1 + m)\tilde{\mathcal{O}}_{\rm a}(\not p_2 - m)\right], \tag{4.145}$$

with

$$\mathcal{O}_{\rm a} = \frac{\not\epsilon_2 \not\epsilon_1 \not k_1}{2m\omega_1}, \qquad \mathcal{O}_{\rm b} = \frac{\not\epsilon_1 \not\epsilon_2 \not k_2}{2m\omega_1}, \tag{4.146}$$

$$\tilde{\mathcal{O}}_{\rm a} = \frac{\not k_1 \not\epsilon_1 \not\epsilon_2}{2m\omega_1}, \qquad \tilde{\mathcal{O}}_{\rm b} = \frac{\not k_2 \not\epsilon_2 \not\epsilon_1}{2m\omega_2}. \tag{4.147}$$

The traces are similar to those encountered in the study of Compton scattering and they can be evaluated following the same procedure. From the results obtained in Compton scattering, one can write down the traces

for pair annihilation into two photons by substituting $\mathbf{k}_1 \to -\mathbf{k}_1$ and $\omega_1 \to -\omega_1$.

$$T_{\rm aa} = -\frac{2}{m\omega_1}\left\{m\omega_2 - 2(-\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2\right\}. \tag{4.148}$$

$$T_{\rm bb} = \frac{2}{m\omega_2}\left\{-m\omega_1 - 2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2\right\}. \tag{4.149}$$

$$T_{\rm ab} = 2\left\{2(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2 - 1\right\} - \frac{2}{m\omega_1}(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2 + \frac{2}{m\omega_2}(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2. \tag{4.150}$$

$$T_{\rm ba} = T_{\rm ab}. \tag{4.151}$$

Density of final states

For the final states of two particles a and b, the density of states ρ_f is given by (*vide* Appendix B)

$$\rho_f = (2\pi)^{-3}E_aE_b\frac{p_a^3 d\Omega_a}{Ep_a^2 - E_a(\boldsymbol{p}\cdot\boldsymbol{p}_a)},$$

where E denotes the total energy and $\boldsymbol{p}$ denotes the total momentum. In the present case for diagram (a),

$$E_a = \omega_1; \quad E_b = \omega_2; \quad E = \omega_1 + \omega_2; \quad \boldsymbol{p} = \boldsymbol{k}_1 + \boldsymbol{k}_2; \quad \boldsymbol{p}_a = \boldsymbol{k}_1; \quad |\boldsymbol{p}_a| = \omega_1.$$

Invoking energy-momentum conservation, we have

$$\begin{aligned}
\mathbf{p}_1 + \mathbf{p}_2 &= \mathbf{k}_1 + \mathbf{k}_2;\\
(\mathbf{p}_1 + \mathbf{p}_2)^2 &= (\mathbf{k}_1 + \mathbf{k}_2)^2;\\
\mathbf{p}_1^2 + \mathbf{p}_2^2 + 2\mathbf{p}_1\cdot\mathbf{p}_2 &= \mathbf{k}^2 + \mathbf{k}_2^2 + \mathbf{k}_1\cdot\mathbf{k}_2;\\
2m^2 + 2mE_2 &= 2\omega_1\omega_2(1-\cos\theta).
\end{aligned} \tag{4.152}$$

Using the above results, we get

$$\begin{aligned}
\rho_f &= \frac{\omega_1\omega_2}{(2\pi)^3}\frac{\omega_1^3 d\Omega_1}{(\omega_1+\omega_2)\omega_1^2 - \omega_1(\omega_1^2 + \boldsymbol{k}_1\cdot\boldsymbol{k}_2)}\\
&= \frac{\omega_1^3\omega_2 d\Omega_1}{(2\pi)^3(\omega_1\omega_2 - \boldsymbol{k}_1\cdot\boldsymbol{k}_2)}.
\end{aligned} \tag{4.153}$$

Using the relation $\boldsymbol{k}_1\cdot\boldsymbol{k}_2 = \omega_1\omega_2\cos\theta$ and Eq. (4.152), we get

$$\rho_f = \frac{\omega_1^3\omega_2 d\Omega_1}{(2\pi)^3 m(m+E_2)}. \tag{4.154}$$

If v is the velocity of the incident positron, then the differential cross section is given by

$$\begin{aligned} d\sigma &= \frac{2\pi}{v}(\Pi_N)|\overline{\mathcal{M}}|^2\rho_f \\ &= \frac{2\pi}{(p_2/E_2)}\frac{1}{2E_1 2E_2 2\omega_1 2\omega_2}|\overline{\mathcal{M}}|^2\frac{\omega_1^3\omega_2 d\Omega_1}{(2\pi)^3 m(m+E_2)} \\ &= \frac{\omega_1^2}{64\pi^2 m^2 p_2(m+E_2)}|\overline{\mathcal{M}}|^2 d\Omega_1. \end{aligned} \tag{4.155}$$

4.6 Bremsstrahlung

Let us consider a moving electron deflected by the Coulomb potential of a nucleus. This will cause an emission of radiation according to the classical theory. This process is known as Bremsstrahlung. We shall calculate the transition probability for the electron to go from one state to another with the emission of a photon in the field of a nucleus. Interaction with an external field is essential for the simultaneous conservation of energy and momentum for the process. In other words, an electron cannot make a transition from one state to another emitting a photon while travelling in a vacuum.

Assuming the Born approximation in which the Coulomb potential acts only once on the electron, there are two modes by which the electron can emit a photon. One mode is that the electron is first scattered by the Coulomb potential and then emits a photon to reach the final state. The other mode is that the electron emits a photon first and then interacts with the Coulomb field. Both these modes are indistinguishable and hence contribute to the total amplitude of transition for the electron from the initial to the final state. The Feynman diagrams for both the modes are depicted in Fig. 4.8.

By an inspection of the two Feynman diagrams (a) and (b), one can write down the total matrix element for the process.

$$\begin{aligned} \mathcal{M} &= -i(4\pi e^2)^{1/2}\left[\bar{u}(\mathbf{p}_2)\not{\epsilon}(\mathbf{k})\frac{1}{\not{p}_1+\not{q}-m}\not{A}(\mathbf{q})u(\mathbf{p}_1)\right. \\ &\qquad \left. + \bar{u}(\mathbf{p}_2)\not{A}(\mathbf{q})\frac{1}{\not{p}_2-\not{q}-m}\not{\epsilon}(\mathbf{k})u(\mathbf{p}_1)\right], \end{aligned} \tag{4.156}$$

with

$$\not{A}(\mathbf{q}) = \frac{4\pi z e^2}{q^2}\delta(\mathbf{q}_0)\gamma_0 \qquad \text{and} \qquad \not{\epsilon} = \sum_\mu \epsilon_\mu \gamma_\mu.$$

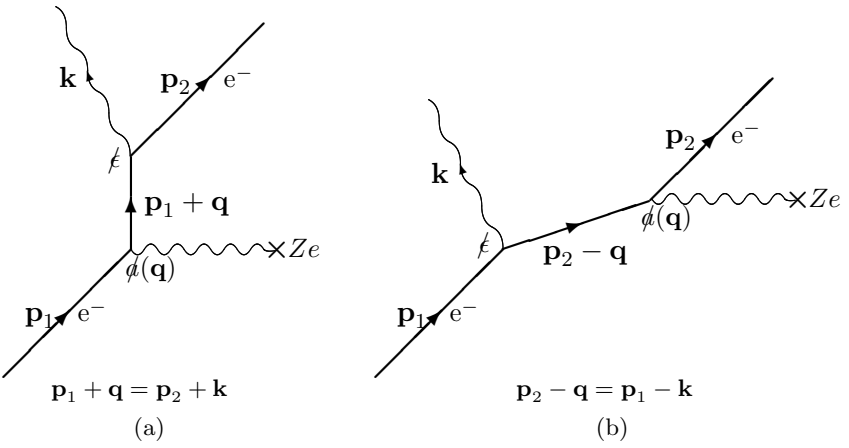


Figure 4.8: Feynman diagrams representing Bremsstrahlung process.

The electron propagators can be rationalized to yield

$$\frac{1}{\not p_1 + \not q - m} = \frac{\not p_1 + \not q + m}{(\not p_1 + \not q)^2 - m^2} = \frac{\not p_1 + \not q + m}{-2\boldsymbol{p}_1 \cdot \boldsymbol{q} - q^2}, \tag{4.157}$$

$$\frac{1}{\not p_2 - \not q - m} = \frac{\not p_2 - \not q + m}{(\not p_2 - \not q)^2 - m^2} = \frac{\not p_2 - \not q + m}{2\boldsymbol{p}_2 \cdot \boldsymbol{q} - q^2}, \tag{4.158}$$

where we have used the condition that in Coulomb scattering, the energy transfer to the electron $q_0 = 0$ and the four-momentum square for the electron $\mathbf{p}_1^2 = m^2$ and $\mathbf{p}_2^2 = m^2$. Using the schematic diagram for the Bremsstrahlung process, the denominators in Eqs. (4.157) and (4.158) can be simplified to yield

$$D_1 = -2\boldsymbol{p}_1 \cdot \boldsymbol{q} - q^2 = 2k(E_2 - p_2 \cos\theta_2); \tag{4.159}$$

$$D_2 = 2\boldsymbol{p}_2 \cdot \boldsymbol{q} - q^2 = -2k(E_1 - p_1 \cos\theta_1); \tag{4.160}$$

where E_1, E_2 denote the energies of the incident and outgoing electron and θ_1, θ_2 denote the angles of the emitted photon and outgoing electron as shown in the schematic diagram 4.9.

Summing over the two spin states of the outgoing electron and the two polarization states of the emitted photon and averaging over the initial spin states of the electron, we can evaluate the differential cross section for the process using the Fermi's golden rule.

$$d\sigma = \frac{2\pi}{v_1} \frac{1}{2E_1 2E_2 2\omega} |\overline{\mathcal{M}}|^2 \rho_f, \tag{4.161}$$

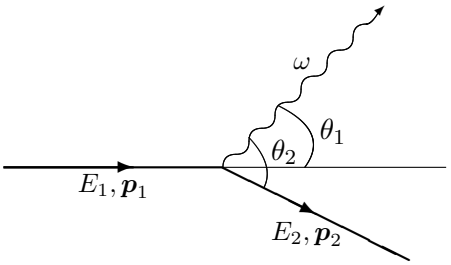


Figure 4.9: Schematic diagram representing the Bremsstrahlung process. The emitted photon makes an angle θ_1 with the incident electron and θ_2 denotes the direction of the outgoing electron. The two planes, one containing incoming electron and outgoing photon and the other containing incoming electron and outgoing electron need not be the same and the angle between them is taken as ϕ. The two planes are not shown explicitly in figure.

where v_1 denotes the energy of the incident electron and ρ_f is the density of final states. The final state consists of the scattered electron, the emitted photon and the nucleus which is infinitely heavy.

$$\rho_f = \frac{1}{(2\pi)^6} E_2 p_2 d\Omega_2 \omega^2 d\omega d\Omega_\omega, \tag{4.162}$$

where E_2, p_2 denote the energy and momentum of the scattered electron, ω, the frequency of the emitted photon which has a continuous spectrum. The solid angles $d\Omega_2$ and $d\Omega_\omega$ correspond to those of scattered electron and emitted photon.

The evaluation of the square of the matrix element is similar to that of Compton scattering and only the final result for the differential cross section is given below[14]:

$$\begin{aligned} d\sigma &= \frac{1}{2\pi}\left(\frac{Ze^2}{q^2}\right)^2 e^2 \frac{d\omega}{\omega}\frac{p_2}{p_1}\sin\theta_2 d\theta_2 \sin\theta_1 d\theta_1 d\phi \\ &\times \left\{ \frac{p_2^2 \sin^2\theta_2 (4E_1^2 - q^2)}{(E_2 - p_2\cos\theta_2)^2} + \frac{p_1^2 \sin^2\theta_1 (4E_2^2 - q^2)}{(E_1 - p_1\cos\theta_1)^2} \right. \\ &\left. - \frac{2Ap_1p_2\sin\theta_1\sin\theta_2\cos\phi - 2B\omega^2}{(E_2 - p_2\cos\theta_2)(E_1 - p_1\cos\theta_1)} \right\}, \end{aligned} \tag{4.163}$$

with

$$A = 4E_1E_2 - q^2 + 2\omega^2, \qquad B = p_2^2\sin^2\theta_2 + p_1^2\sin^2\theta_1.$$

[14]R. P. Feynman, *Quantum Electrodynamics* W. A. Benjamin (1976).

4.6.1 In the soft photon limit

If the emitted photon energy is small compared to the rest mass of the electron (in the soft photon limit), then an approximate expression for the Bremsstrahlung cross section which has a simple interpretation can be obtained.

Writing the matrix element in terms of $\not{k}$ instead of $\not{q}$, we have

$$\begin{aligned}\mathcal{M} &= -i(4\pi e^2)^{1/2}\left[\bar{u}(\mathbf{p}_2)\not{\epsilon}(\mathbf{k})\frac{1}{\not{p}_2+\not{k}-m}\not{\epsilon}(\mathbf{q})u(\mathbf{p}_1)\right.\\ &\qquad\left.+\,\bar{u}(\mathbf{p}_2)\not{\epsilon}(\mathbf{q})\frac{1}{\not{p}_1-\not{k}-m}\not{\epsilon}(\mathbf{k})u(\mathbf{p}_1)\right], \quad (4.164)\end{aligned}$$

Rationalizing the denominator of the matrix element, we get

$$\begin{aligned}\mathcal{M} &= -i(4\pi e^2)^{1/2}\frac{4\pi z e^2}{q^2}\bar{u}(\mathbf{p}_2)\left[\not{\epsilon}(\mathbf{k})\frac{\not{p}_2+\not{k}+m}{(\not{p}_2+\not{k})^2-m^2}\gamma_0\right.\\ &\qquad\left.+\,\gamma_0\frac{\not{p}_1-\not{k}+m}{(\not{p}_1-\not{k})^2-m^2}\not{\epsilon}(\mathbf{k})\right]u(\mathbf{p}_1)\\ &= -i(4\pi e^2)^{1/2}\frac{4\pi z e^2}{q^2}\bar{u}(\mathbf{p}_2)\left[\not{\epsilon}(\mathbf{k})\frac{\not{p}_2+\not{k}+m}{2\mathbf{p}_2\cdot\mathbf{k}}\gamma_0\right.\\ &\qquad\left.+\,\gamma_0\frac{\not{p}_1-\not{k}+m}{(-2\mathbf{p}_1\cdot\mathbf{k})}\not{\epsilon}(\mathbf{k})\right]u(\mathbf{p}_1), \quad (4.165)\end{aligned}$$

since

$$\begin{aligned}(\not{p}_2+\not{k})^2-m^2 &= \mathbf{p}_2^2+\mathbf{k}^2+2\mathbf{p}_2\cdot\mathbf{k}-m^2=2\mathbf{p}_2\cdot\mathbf{k}, \quad (4.166)\\ (\not{p}_1-\not{k})^2-m^2 &= \mathbf{p}_1^2+\mathbf{k}^2-2\mathbf{p}_1\cdot\mathbf{k}-m^2=-2\mathbf{p}_1\cdot\mathbf{k}. \quad (4.167)\end{aligned}$$

It is convenient to write the matrix element (4.165) in a simple form.

$$\mathcal{M} = -i(4\pi e^2)^{1/2}\frac{4\pi z e^2}{q^2}\left\{\frac{T_a}{2\mathbf{p}_2\cdot\mathbf{k}}-\frac{T_b}{2\mathbf{p}_1\cdot\mathbf{k}}\right\}, \quad (4.168)$$

with

$$\begin{aligned}T_a &= \bar{u}(\mathbf{p}_2)\not{\epsilon}(\mathbf{k})(\not{p}_2+\not{k}+m)\gamma_0 u(\mathbf{p}_1) \quad (4.169)\\ T_b &= \bar{u}(\mathbf{p}_2)\gamma_0(\not{p}_1-\not{k}+m)\not{\epsilon}(\mathbf{k})u(\mathbf{p}_1) \quad (4.170)\end{aligned}$$

In the soft photon limit, the $\not{k}$ term in T_a and T_b can be neglected. Then

$$\begin{aligned}T_a &= \bar{u}(\mathbf{p}_2)\left\{\not{\epsilon}(\mathbf{k})(\not{p}_2+m)\gamma_0\right\}u(\mathbf{p}_1)\\ &= \bar{u}(\mathbf{p}_2)\left(-\not{p}_2\not{\epsilon}(\mathbf{k})+2\mathbf{p}_2\cdot\boldsymbol{\varepsilon}+m\not{\epsilon}(\mathbf{k})\right)\gamma_0 u(\mathbf{p}_1)\\ &= \bar{u}(\mathbf{p}_2)\gamma_0 u(\mathbf{p}_1)2\mathbf{p}_2\cdot\boldsymbol{\varepsilon}, \quad \text{since } \bar{u}(\mathbf{p}_2)\not{p}_2=m\bar{u}(\mathbf{p}_2). \quad (4.171)\end{aligned}$$

In a similar way, we get

$$\begin{aligned} T_b &= \bar{u}(\mathbf{p}_2)\left\{\gamma_0(\not{p}_1 + m)\not{\varepsilon}(\mathbf{k})\right\}u(\mathbf{p}_1) \\ &= \bar{u}(\mathbf{p}_2)\gamma_0\left(-\not{\varepsilon}(\mathbf{k})\not{p}_1 + 2\mathbf{p}_1\cdot\boldsymbol{\varepsilon} + m\not{\varepsilon}(\mathbf{k})\right)u(\mathbf{p}_1) \\ &= \bar{u}(\mathbf{p}_2)\gamma_0 u(\mathbf{p}_1)2\mathbf{p}_1\cdot\boldsymbol{\varepsilon}, \quad \text{since } \not{p}_1 u(\mathbf{p}_1) = mu(\mathbf{p}_1). \quad (4.172) \end{aligned}$$

Substituting these results in Eq. (4.168), we get

$$\mathcal{M} = -i(4\pi e^2)^{1/2}\frac{4\pi z e^2}{q^2}\left(\bar{u}(\mathbf{p}_2)\gamma_0 u(\mathbf{p}_1)\right)\left(\frac{\mathbf{p}_2\cdot\boldsymbol{\varepsilon}(\mathbf{k})}{\mathbf{p}_2\cdot\mathbf{k}} - \frac{\mathbf{p}_1\cdot\boldsymbol{\varepsilon}(\mathbf{k})}{\mathbf{p}_1\cdot\mathbf{k}}\right) \quad (4.173)$$

Squaring the matrix element and taking the sum over the final spin states and the average over the initial spin states of the electron, we get

$$\begin{aligned} d\sigma &= \frac{2\pi}{v_1}\frac{1}{2E_1 2E_2 2\omega}|\overline{\mathcal{M}}|^2\frac{E_2 p_2 d\Omega_2}{(2\pi)^3}\frac{\omega^2 d\omega d\Omega_\omega}{(2\pi)^3} \\ &= \left[\frac{2\pi}{v_1}\frac{1}{2E_1 2E_2}\left(\frac{4\pi z e^2}{q^2}\right)^2\sum_{\text{spins}}|\bar{u}(\mathbf{p}_2)\gamma_0 u(\mathbf{p}_1)|^2\frac{E_2 p_2 d\Omega_2}{(2\pi)^3}\right] \\ &\quad\times\left[\frac{(4\pi e^2)\omega^2 d\Omega_\omega}{2\omega(2\pi)^3}\left(\frac{\mathbf{p}_2\cdot\boldsymbol{\varepsilon}(\mathbf{k})}{\mathbf{p}_2\cdot\mathbf{k}} - \frac{\mathbf{p}_1\cdot\boldsymbol{\varepsilon}(\mathbf{k})}{\mathbf{p}_1\cdot\mathbf{k}}\right)^2\right], \quad (4.174) \end{aligned}$$

where the first square bracket denotes the elastic scattering cross section of electron in the Coulomb field making a transition from an initial state 1 to a final state 2 and the second square bracket gives the probability of emission of photon in frequency range $d\omega$ and solid angle $d\Omega_\omega$.

4.7 Electron-positron Pair Production

A single photon of energy greater than twice the electron mass cannot create an electron-positron pair since energy and momentum has to be conserved at each vertex and also for the whole process. Although two photons can create a pair, this process is extremely unlikely since the photon density is generally very low. However, a single photon can create a pair in the Coulomb field of a nucleus by sharing some momentum as shown in Fig. 4.10. It can happen in two ways: (a) The produced positron (virtual) can share momentum with the external field and emerge as a real positron or (b) the produced electron (virtual) can share momentum with the external field and emerge as real electron. These two modes are

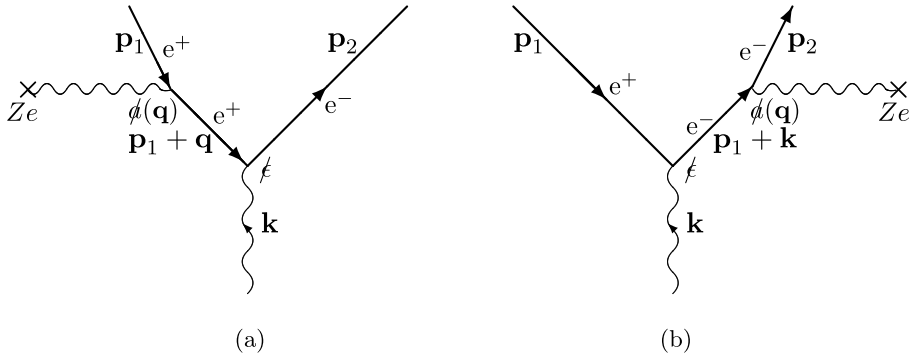


Figure 4.10: Electron-positron pair production in the Coulomb field of a nucleus by a single photon. The four-momenta satisfy the relation $\mathbf{p}_1 + \mathbf{q} + \mathbf{k} = \mathbf{p}_2$.

indistinguishable and so the two amplitudes should be added to compute the cross section for the process.

By an inspection of the two Feynman diagrams (a) and (b), one can write down the total matrix element for the process.

$$\begin{aligned}\mathcal{M} &= -i(4\pi e^2)^{1/2}\left[u(\mathbf{p}_2)\not{\epsilon}(\mathbf{k})\frac{1}{\not{p}_1 + \not{q} + m}\not{\phi}(\mathbf{q})v(\mathbf{p}_1)\right. \\ &\quad \left. + u(\mathbf{p}_2)\not{\phi}(\mathbf{q})\frac{1}{\not{p}_1 + \not{k} - m}\not{\epsilon}(\mathbf{k})v(\mathbf{p}_1)\right], \quad (4.175)\end{aligned}$$

with

$$\not{\phi}(\mathbf{q}) = \frac{4\pi z e^2}{q^2}\delta(\mathbf{q}_0)\gamma_0 \qquad \text{and} \qquad \not{\epsilon} = \sum_\mu \epsilon_\mu \gamma_\mu.$$

The matrix elements for this process are similar to those of Bremsstrahlung and can be evaluated in a similar way.

4.8 Muon Pair Production in Electron-positron Collision

Hitherto, we have considered only processes involving electrons, positrons and photons. The formalism can easily be extended to include production of new particles in $e^- - e^+$ collisions.

Let us consider the process

$$e^- + e^+ \longrightarrow \mu^- + \mu^+.$$

This process is possible if the energy of the colliding partners e^- and e^+ in the c.m. frame is greater than the mass equivalent of μ^- and μ^+ *i.e.* 212 MeV. At such energies, one can neglect the electron mass (0.511 MeV) in the calculation.

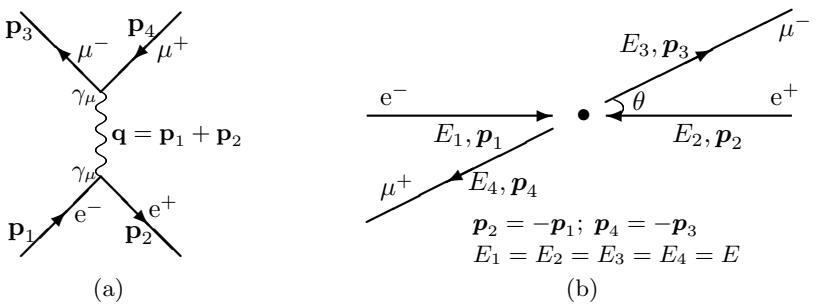


Figure 4.11: Muon production in e^-, e^+ collision. The Feynman diagram (a) represents the process $e^-e^+ \longrightarrow \mu^-\mu^+$. Diagram (b) represents the kinematics for the process in c.m. frame. $E_1, \boldsymbol{p}_1$ and $E_2, \boldsymbol{p}_2$ denote the energy and momentum of the colliding partners e^- and e^+ and $E_3, \boldsymbol{p}_3$ and $E_4, \boldsymbol{p}_4$ denote the energy and momentum of the produced particles μ^- and μ^+. θ denotes the angle at which the produced particles emerge in the c.m. system.

Using the Feynman rules, the matrix element for the process $e^-e^+ \longrightarrow \mu^-\mu^+$ can be written down. For writing down the matrix elements, we follow the flow direction of arrow marks in Feynman diagram. We follow the same convention that we used earlier and write down the spinors $u(\mathbf{p}_1), \bar{v}(\mathbf{x}_2), \bar{u}(\mathbf{x}_3), (\bar{v}(\mathbf{x}_4)$ as $u_1, \bar{v}_2, \bar{u}_3, \bar{v}_4$.

$$\mathcal{M} = i4\pi e^2 \frac{[(\bar{v}_2\gamma_\mu u_1)(\bar{u}_3\gamma_\mu v_4)]}{(\mathbf{p}_1 + \mathbf{p}_2)^2}, \tag{4.176}$$

where u_1 and u_3 denote the spinors corresponding to particles (e^- and μ^-) and v_2 and v_4 denote the spinors corresponding to antiparticles (e^+ and μ^+). The four momenta of the colliding electron and positron are denoted by $\mathbf{p}_1$ and $\mathbf{p}_2$ and the four-momenta of the produced μ^- and μ^+ are denoted by $\mathbf{p}_3$ and $\mathbf{p}_4$. The transition probability is the absolute square of the matrix element (4.176). Since the spins are not observed, a sum over the final spin states and an average over the initial spin states

have to be taken. This introduces a factor $\frac{1}{4}$.

$$\begin{aligned}
|\overline{\mathcal{M}}|^2 &= \frac{1}{4}\mathcal{M}\mathcal{M}^\dagger \\
&= \frac{4\pi^2 e^4}{(\mathbf{p}_1+\mathbf{p}_2)^4}\sum_{\text{spins}}\left[(\bar{v}_2\gamma_\mu u_1)(\bar{u}_3\gamma_\mu v_4)\right]\left[(\bar{v}_2\gamma_\nu u_1)(\bar{u}_3\gamma_\nu v_4)\right]^\dagger \\
&= \frac{4\pi^2 e^4}{(\mathbf{p}_1+\mathbf{p}_2)^4}\sum_{\text{spins}}\left[(\bar{v}_2\gamma_\mu u_1)(\bar{v}_2\gamma_\nu u_1)^\dagger\right]\left[(\bar{u}_3\gamma_\mu v_4)(\bar{u}_3\gamma_\nu v_4)^\dagger\right]. \\
&= \frac{4\pi^2 e^4}{(\mathbf{p}_1+\mathbf{p}_2)^4}\sum_{\text{spins}}\left[(\bar{v}_2\gamma_\mu u_1)(\bar{u}_1\gamma_\nu v_2)\right]\left[(\bar{u}_3\gamma_\mu v_4)(\bar{v}_4\gamma_\nu u_3)\right]. \quad (4.177)
\end{aligned}$$

Using the trace techniques, we obtain

$$\begin{aligned}
|\overline{\mathcal{M}}|^2 &= \frac{4\pi^2 e^4}{(\mathbf{p}_1+\mathbf{p}_2)^4}\text{Tr}[\gamma_\mu(\not{p}_1+m_e)\gamma_\nu(\not{p}_2-m_e)] \\
&\quad\times\text{Tr}[\gamma_\mu(\not{p}_4-m_\mu)\gamma_\nu(\not{p}_3+m_\mu)], \quad (4.178)
\end{aligned}$$

where m_e and m_μ denote the mass of the electron and muon respectively. Let us now evaluate the traces, neglecting the terms involving m_e, since the energy of the colliding e^- and e^+ should be greater than twice the muon mass (~ 212 MeV). Since the trace of a product of odd number of γ matrices vanishes, we get

$$\begin{aligned}
\text{Tr}[\gamma_\mu(\not{p}_1+m_e)\gamma_\nu(\not{p}_2-m_e)] &= \text{Tr}[\gamma_\mu\not{p}_1\gamma_\nu\not{p}_2] \\
&= 4p_{1\mu}p_{2\nu}+4p_{2\mu}p_{1\nu}-4(\mathbf{p}_1\cdot\mathbf{p}_2)g_{\mu\nu}; \quad (4.179) \\
\text{Tr}[\gamma_\mu(\not{p}_4-m_\mu)\gamma_\nu(\not{p}_3+m_\mu)] &= \text{Tr}[\gamma_\mu\not{p}_4\gamma_\nu\not{p}_3]-m_\mu^2\text{Tr}[\gamma_\mu\gamma_\nu] \\
&= 4p_{3\mu}p_{4\nu}+4p_{4\mu}p_{3\nu}-4(\mathbf{p}_3\cdot\mathbf{p}_4)g_{\mu\nu} \\
&\quad -4m_\mu^2\, g_{\mu\nu}. \quad (4.180)
\end{aligned}$$

Inserting (4.179) and (4.180) in Eq. (4.178), we get

$$\begin{aligned}
|\overline{\mathcal{M}}|^2 &= \frac{4\pi^2 e^4}{(\mathbf{p}_1+\mathbf{p}_2)^4}[4p_{1\mu}p_{2\nu}+4p_{2\mu}p_{1\nu}-4(\mathbf{p}_1\cdot\mathbf{p}_2)g_{\mu\nu}] \\
&\quad\times[4p_{3\mu}p_{4\nu}+4p_{4\mu}p_{3\nu}-4(\mathbf{p}_3\cdot\mathbf{p}_4)g_{\mu\nu}-4m_\mu^2\, g_{\mu\nu}] \\
&= \frac{128\pi^2 e^4}{(\mathbf{p}_1+\mathbf{p}_2)^4}[(\mathbf{p}_1\cdot\mathbf{p}_3)(\mathbf{p}_2\cdot\mathbf{p}_4)+(\mathbf{p}_1\cdot\mathbf{p}_4)(\mathbf{p}_2\cdot\mathbf{p}_3)+m_\mu^2(\mathbf{p}_1\cdot\mathbf{p}_2)]. \\
&= \frac{128\pi^2 e^4}{(\mathbf{p}_1+\mathbf{p}_2)^4}[(\mathbf{p}_1\cdot\mathbf{p}_3)^2+(\mathbf{p}_1\cdot\mathbf{p}_4)^2+m_\mu^2(\mathbf{p}_1\cdot\mathbf{p}_2)]. \quad (4.181)
\end{aligned}$$

The last step in Eq. (4.181) is obtained using the energy-momentum conservation in the reaction

$$\mathbf{p}_1 + \mathbf{p}_2 = \mathbf{p}_3 + \mathbf{p}_4,$$

which can be equally expressed as

$$\mathbf{p}_1 - \mathbf{p}_3 = \mathbf{p}_4 - \mathbf{p}_2 \longrightarrow \mathbf{p}_1 \cdot \mathbf{p}_3 = \mathbf{p}_2 \cdot \mathbf{p}_4.$$

$$\mathbf{p}_1 - \mathbf{p}_4 = \mathbf{p}_3 - \mathbf{p}_2 \longrightarrow \mathbf{p}_1 \cdot \mathbf{p}_4 = \mathbf{p}_2 \cdot \mathbf{p}_3.$$

Let us evaluate $|\overline{\mathcal{M}}|^2$ given by Eq. (4.181) in c.m. frame using the kinematics described in Fig. (b).

$$\begin{aligned}
\mathbf{p}_1 \cdot \mathbf{p}_3 &= E_1 E_3 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_3 = E^2 - Ep\cos\theta; && (4.182)\\
\mathbf{p}_1 \cdot \mathbf{p}_4 &= E_1 E_4 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_4 = E^2 + Ep\cos\theta; && (4.183)\\
\mathbf{p}_1 \cdot \mathbf{p}_2 &= E_1 E_2 - \boldsymbol{p}_1 \cdot \boldsymbol{p}_2 = 2E^2; && (4.184)\\
(\mathbf{p}_1 + \mathbf{p}_2)^2 &= \mathbf{p}_1^2 + \mathbf{p}_2^2 + 2\mathbf{p}_1 \cdot \mathbf{p}_2 \approx 4E^2; && (4.185)
\end{aligned}$$

since

$$\begin{aligned}
&E_1 = E_2 = E_3 = E_4 = E; \;\; \boldsymbol{p}_2 = -\boldsymbol{p}_1; \;\; \boldsymbol{p}_4 = -\boldsymbol{p}_3;\\
&|\boldsymbol{p}_1| = |\boldsymbol{p}_2| = E; \;\; |\boldsymbol{p}_3| = |\boldsymbol{p}_4| = p = (E^2 - m_\mu^2)^{1/2};\\
&\mathbf{p}_1^2 = \mathbf{p}_2^2 = m_e^2, \quad \text{negligible when compared to their energy } E.
\end{aligned}$$

Substituting (4.182) - (4.185) into Eq. (4.181) and simplifying, we get

$$\begin{aligned}
|\overline{\mathcal{M}}|^2 &= \frac{8\pi^2 e^4}{E^4}\left[(E^2 - Ep\,\cos\theta)^2 + (E^2 + Ep\,\cos\theta)^2 + 2E^2 m_\mu^2\right]\\
&= 16\pi^2 e^4 \left[(1+\cos^2\theta) + \frac{m_\mu^2}{E^2}(1-\cos^2\theta)\right]. \qquad (4.186)
\end{aligned}$$

The differential cross section is given by Fermi's golden rule

$$d\sigma = \frac{2\pi}{v}(\Pi N)^{-1}|\overline{\mathcal{M}}|^2 \rho_f,$$

with

$$v = 2\frac{p_1}{E_1} \approx 2; \;\; (\Pi N)^{-1} = \frac{1}{2E_1 2E_2 2E_3 2E_4} = \frac{1}{16E^4}; \;\; \rho_f = \frac{E_3 p_3}{2(2\pi)^3} d\Omega_3.$$

Hence

$$\frac{d\sigma}{d\Omega_3} = \frac{e^4}{16E^2}\left(1-\frac{m_\mu^2}{E^2}\right)^{1/2}\left[1+\cos^2\theta+\frac{m_\mu^2}{E^2}(1-\cos^2\theta)\right]. \quad (4.187)$$

Integrating over $d\Omega_3 = 2\pi\sin\theta d\theta$, we get

$$\int(1+\cos^2\theta)d\Omega = \frac{16}{3}\pi; \qquad \int(1-\cos^2\theta)d\Omega = \frac{8}{3}\pi;$$

which, in turn, yields the total cross section.

$$\begin{aligned}\sigma &= \frac{\pi e^4}{3E^2}\left(1-\frac{m_\mu^2}{E^2}\right)^{1/2}\left(1+\frac{m_\mu^2}{2E^2}\right)\\ &= \frac{\pi\alpha^2}{3E^2}\left(1-\frac{m_\mu^2}{E^2}\right)^{1/2}\left(1+\frac{m_\mu^2}{2E^2}\right), \end{aligned} \quad (4.188)$$

where $\alpha = e^2$ denotes the fine structure constant.

4.8.1 Application to hadron production

We have already calculated the cross section for the elastic scattering of e^-e^+. If the c.m. energy is sufficiently high, the inelastic events begin to contribute and if the energy exceeds several GeV, then several hadrons can also be produced. Thus the e^-e^+ collision at high energies serve as a convenient tool for searching for new particles.

In the GeV range, one can neglect the muon mass and the total cross section for $e^-e^+ \to \mu^-\mu^+$ takes a simple form

$$\sigma = \frac{\pi\alpha^2}{3E^2} = \frac{4\pi\alpha^2}{3s^2}, \quad (4.189)$$

where $s = 4E^2$ denotes the square of total energy in c.m. system.

Equation (4.189) has been used successfully to determine the number of colour degree of freedom for the quarks. Since the hadrons are made up of quarks, the hadronic final state in e^-e^+ collision occurs through the events $e^-e^+ \to q_i\bar{q}_i$, where $q_i, \bar{q}_i$ denote the quark-antiquark pair of flavour i. If e_i is the fractional charge of the quark q_i, then the production cross section for $e^-e^+ \to q_i\bar{q}_i$ is

$$\sigma = \frac{\pi\alpha^2}{3E^2}N_c e_i^2, \quad (4.190)$$

where e_i denotes the fractional charge of quark of flavour i and N_c denotes the number of colours. The ratio

$$R = \frac{\sigma(e^-e^+ \to hadrons)}{\sigma(e^-e^+ \to \mu^-\mu^+)} = N_c \sum_i e_i^2, \tag{4.191}$$

Restricting the quark flavours to u, d, s, c and choosing $N_c = 3$, the ratio of the cross sections R is given in the following table.

Table 4.1: Ratio of cross sections R calculated with increasing c.m. collision energies indicating opening of new channels. Quark flavours: up(u), down(d), strange(s), charm(c). $N_c = 3$.

Quark flavour (i)	charge (e_i)	$\sum_i e_i^2$	$R = N_c \sum_i e_i^2$
u, d	$\frac{2}{3}, -\frac{1}{3}$	$(\frac{2}{3})^2 + (-\frac{1}{3})^2 = \frac{5}{9}$	$\frac{5}{3}$
u, d, s	$\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}$	$\frac{2}{3}$	2
u, d, s, c	$\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}, \frac{2}{3}$	$\frac{10}{9}$	$\frac{10}{3}$

The (e^-e^+) colliding beam experiments performed in the energy range 2.5 GeV - 40 GeV confirm the above theoretical prediction for R and this is taken as the sufficient evidence for the three colour degrees of freedom for quarks. The sudden increase in the ratio of R at certain energies points to the opening of new channels in the production of hadrons.

Review Questions

4.1 Draw the Feynman diagram for the Coulomb scattering of an electron by a nucleus of charge Ze and deduce the Mott scattering formula. Under what condition, it will reduce to the Rutherford scattering formula.

4.2 Draw the lowest order Feynman diagrams for the Compton scattering and write down the Feynman amplitudes for them. Deduce the Klein-Nishina formula. Show, how it reduces to the Thomson scattering formula in the non-relativistic limit.

4.3 Explain how you get two Feynman diagrams in the study of electron-electron scattering by one-photon exchange. Write down the Feynman diagrams and discuss how you can obtain the cross section for the process.

4.4 Draw the Feynman diagrams for electron-positron scattering and explain their significance. Discuss how you can obtain the differential cross section for this process.

4.5 Discuss the process of electron-positron pair annihilation into two photons and explain how you can determine the cross section for this process.

4.6 Discuss the process of emission of photons known as bremsstrahlung during the Coulomb scattering of an electron. For emission of soft photons, show the cross section can be expressed as a product of two factors, one giving the probability of Coulomb scattering and the other giving the probability of photon emission in a frequency interval.

4.7 Consider the process $e^-e^+ \longrightarrow \mu^-\mu^+$ and show that the total cross section for this process at very high energies is given by a simple expression $\sigma = \frac{4\pi\alpha^2}{3s}$, where α denotes the fine structure constant and s denotes the total c.m. energy of colliding particles.

Problems

4.1 Using the anticommutation relations of γ matrices, prove the following formulae involving a product of γ matrices of repeated index[15]:

$$\gamma_\mu\gamma_\mu = 4; \qquad \gamma_\mu\gamma_\nu\gamma_\mu = -2\gamma_\nu;$$
$$\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\mu = 4g_{\nu\rho}; \qquad \gamma_\mu\gamma_\nu\gamma_\rho\gamma_\lambda\gamma_\mu = -2\gamma_\lambda\gamma_\rho\gamma_\nu.$$

4.2 In the discussion on Compton scattering, we have come across the following traces as given by Eqs. (4.54) - (4.57):

$$\begin{aligned}
T_{\rm aa} &= {\rm Tr}\left[\mathcal{O}_{\rm a}(\not{p}_1 + m)\tilde{\mathcal{O}}_{\rm a}(\not{p}_2 + m)\right];\\
T_{\rm ab} &= {\rm Tr}\left[\mathcal{O}_{\rm a}(\not{p}_1 + m)\tilde{\mathcal{O}}_{\rm b}(\not{p}_2 + m)\right];\\
T_{\rm ba} &= {\rm Tr}\left[\mathcal{O}_{\rm b}(\not{p}_1 + m)\tilde{\mathcal{O}}_{\rm a}(\not{p}_2 + m)\right];\\
T_{\rm bb} &= {\rm Tr}\left[\mathcal{O}_{\rm b}(\not{p}_1 + m)\tilde{\mathcal{O}}_{\rm b}(\not{p}_2 + m)\right];
\end{aligned}$$

with the operators, explicitly given by Eqs. (4.58) - (4.61).

$$\begin{aligned}
\mathcal{O}_{\rm a} &= \frac{1}{2m\omega_1}\not{\epsilon}_2\not{\epsilon}_1\not{k}_1,\\
\mathcal{O}_{\rm b} &= \frac{1}{2m\omega_2}\not{\epsilon}_1\not{\epsilon}_2\not{k}_2,\\
\tilde{\mathcal{O}}_{\rm a} &= \frac{1}{2m\omega_1}\gamma_0(\not{\epsilon}_2\not{\epsilon}_1\not{k}_1)^\dagger\gamma_0 = \frac{1}{2m\omega_1}(\not{k}_1\not{\epsilon}_1\not{\epsilon}_2),\\
\tilde{\mathcal{O}}_{\rm b} &= \frac{1}{2m\omega_2}\gamma_0(\not{\epsilon}_1\not{\epsilon}_2\not{k}_2)^\dagger\gamma_0 = \frac{1}{2m\omega_2}(\not{k}_2\not{\epsilon}_2\not{\epsilon}_1),
\end{aligned}$$

[15] A summation over the repeated index is implied.

Evaluate them and obtain the results given by Eqs. (4.62) - (4.65).

$$\begin{aligned} T_{\rm aa} &= \frac{2}{m\omega_1}\left\{m\omega_2 + 2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2\right\}. \\ T_{\rm bb} &= \frac{2}{m\omega_2}\left\{m\omega_1 - 2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2\right\}. \\ T_{\rm ab} &= 2\left\{2(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2 - 1\right\} - \frac{2}{m\omega_1}(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2 + \frac{2}{m\omega_2}(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2. \\ T_{\rm ba} &= T_{\rm ab}. \end{aligned}$$

4.3 Given the Klein-Nishina formula (4.73), obtain the total cross section for Compton scattering by performing the angular integration.

4.4 Draw the Feynman diagrams for positron-positron scattering and show that the differential cross section for this process is the same as that of electron-electron scattering..

Solutions to Problems

4.1

$$\gamma_\mu\gamma_\mu = \gamma_0^2 - \gamma_x^2 - \gamma_y^2 - \gamma_z^2 = 4.$$

Using the relation $\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}$ repeatedly, we get[16]

$$\begin{aligned} \gamma_\mu \underbrace{\gamma_\nu\gamma_\mu} &= \gamma_\mu(-\gamma_\mu\gamma_\nu + 2g_{\mu\nu}) = -4\gamma_\nu + 2\gamma_\nu = -2\gamma_\nu \\ \gamma_\mu\gamma_\nu \underbrace{\gamma_\rho\gamma_\mu} &= \gamma_\mu\gamma_\nu(-\gamma_\mu\gamma_\rho + 2g_{\mu\rho}) \\ &= -\gamma_\mu \underbrace{\gamma_\nu\gamma_\mu}\gamma_\rho + 2g_{\mu\rho}\gamma_\mu\gamma_\nu \\ &= \gamma_\mu\gamma_\mu\gamma_\nu\gamma_\rho - 2g_{\mu\nu}\gamma_\mu\gamma_\rho + 2g_{\mu\rho}\gamma_\mu\gamma_\nu \\ &= 4\gamma_\nu\gamma_\rho - 2\gamma_\nu\gamma_\rho + 2\gamma_\rho\gamma_\nu \\ &= 2\gamma_\nu\gamma_\rho + 2\gamma_\rho\gamma_\nu = 4g_{\nu\rho}. \end{aligned}$$

$$\begin{aligned} \gamma_\mu\gamma_\nu\gamma_\rho\underbrace{\gamma_\lambda\gamma_\mu} &= \gamma_\mu\gamma_\nu\gamma_\rho(-\gamma_\mu\gamma_\lambda + 2g_{\mu\lambda}) \\ &= -\gamma_\mu\gamma_\nu\underbrace{\gamma_\rho\gamma_\mu}\gamma_\lambda + 2g_{\mu\lambda}\gamma_\mu\gamma_\nu\gamma_\rho \\ &= \gamma_\mu\underbrace{\gamma_\nu\gamma_\mu}\gamma_\rho\gamma_\lambda - 2g_{\rho\mu}\gamma_\mu\gamma_\nu\gamma_\lambda + 2g_{\mu\lambda}\gamma_\mu\gamma_\nu\gamma_\rho \\ &= -\gamma_\mu\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\lambda + 2g_{\mu\nu}\gamma_\mu\gamma_\rho\gamma_\lambda - 2g_{\rho\mu}\gamma_\mu\gamma_\nu\gamma_\lambda + 2g_{\mu\lambda}\gamma_\mu\gamma_\nu\gamma_\rho \\ &= -4\gamma_\nu\gamma_\rho\gamma_\lambda + 2\gamma_\nu\gamma_\rho\gamma_\lambda - 2\gamma_\rho\gamma_\nu\gamma_\lambda + 2\gamma_\lambda\gamma_\nu\gamma_\rho \\ &= -2\gamma_\nu\gamma_\rho\gamma_\lambda - 2\gamma_\rho\gamma_\nu\gamma_\lambda + 2\gamma_\lambda\gamma_\nu\gamma_\rho \\ &= -4g_{\nu\rho}\gamma_\lambda + 2\gamma_\lambda\gamma_\nu\gamma_\rho = -2\gamma_\lambda\gamma_\rho\gamma_\nu. \end{aligned}$$

The last step is obtained using the relation $\gamma_\nu\gamma_\rho = -\gamma_\rho\gamma_\nu + 2g_{\nu\rho}$.

[16] Underbraces are used to highlight the terms that are used in the step-by-step calculation.

4.2 To derive $T_{\rm aa}$

The following general relations are used in the evaluation of the traces.[17]

$$\not{a}\not{b} = -\not{b}\not{a} + 2\mathbf{a}\cdot\mathbf{b}, \qquad (\mathbf{a}\cdot\mathbf{b} = a_\mu b_\mu = a_0 b_0 - \boldsymbol{a}\cdot\boldsymbol{b}); \tag{4.192}$$

$$\not{a}\not{a} = \mathbf{a}\cdot\mathbf{a}; \tag{4.193}$$

$$\mathrm{Tr}(\not{a}\not{b}) = 4\mathbf{a}\cdot\mathbf{b}; \tag{4.194}$$

$$\mathrm{Tr}(\not{a}\not{b}\not{c}\not{d}) = 4(\mathbf{a}\cdot\mathbf{b})(\mathbf{c}\cdot\mathbf{d}) - 4(\mathbf{a}\cdot\mathbf{c})(\mathbf{b}\cdot\mathbf{d}) + 4(\mathbf{a}\cdot\mathbf{d})(\mathbf{b}\cdot\mathbf{c}); \tag{4.195}$$

where $\mathbf{a}$ and $\mathbf{b}$ are any four vectors. The trace of a product of an odd number of γ matrices vanishes. The trace of a product of γ matrices remains the same if the product is taken in the reverse order.

The trace of a product of matrices A, B, C is invariant under cyclic permutations and it obeys the distribution law.

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) = \mathrm{Tr}(CAB); \tag{4.196}$$

$$\mathrm{Tr}(A+B+C) = \mathrm{Tr}A + \mathrm{Tr}B + \mathrm{Tr}C. \tag{4.197}$$

In the present case,

$$\not{p}_1\not{p}_1 = \mathbf{p}_1\cdot\mathbf{p}_1 = m^2; \qquad \not{p}_2\not{p}_2 = \mathbf{p}_2\cdot\mathbf{p}_2 = m^2. \tag{4.198}$$

$$\not{k}_1\not{k}1 = \mathbf{k}_1\cdot\mathbf{k}_1 = 0; \qquad \not{k}_2\not{k}_2 = \mathbf{k}_2\cdot\mathbf{k}_2 = 0. \tag{4.199}$$

$$\not{k}_1\not{\varepsilon}_1 = -\not{\varepsilon}_1\not{k}_1 + 2\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_1 = -\not{\varepsilon}_1\not{k}_1; \tag{4.200}$$

$$\not{k}_2\not{\varepsilon}_2 = -\not{\varepsilon}_2\not{k}_2 + 2\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_2 = -\not{\varepsilon}_2\not{k}_2; \tag{4.201}$$

since $\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_1 = 0$ and $\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_2 = 0$ because $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ have no time components but their space components are normal to the momentum vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ respectively. Also

$$\not{\varepsilon}_1\not{\varepsilon}_1 = \boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_1 = -1; \qquad \not{\varepsilon}_2\not{\varepsilon}_2 = \boldsymbol{\varepsilon}_2\cdot\boldsymbol{\varepsilon}_2 = -1. \tag{4.202}$$

In addition, if we choose the coordinate system to coincide with the rest system of the initial electron, then

$$\not{p}_1\not{\varepsilon}_1 = -\not{\varepsilon}_1\not{p}_1 + 2\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_1 = -\not{\varepsilon}_1\not{p}_1, \tag{4.203}$$

$$\not{p}_1\not{\varepsilon}_2 = -\not{\varepsilon}_2\not{p}_1 + 2\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_2 = -\not{\varepsilon}_2\not{p}_1, \tag{4.204}$$

since $\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_1 = \mathbf{p}_1\cdot\boldsymbol{\varepsilon}_2 = 0$ because $\mathbf{p}_1$ has only the time component whereas $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ have only the space components.

$$T_{\rm aa} = \mathrm{Tr}\left[\mathcal{O}_{\rm a}(\not{p}_1 + m)\tilde{\mathcal{O}}_{\rm a}(\not{p}_2 + m)\right], \tag{4.205}$$

[17] We use bold letters $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ to denote the four-vectors and bold italic letters $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}$ to denote the three vectors.

where

$$\begin{aligned}\mathcal{O}_{\rm a} &= \frac{1}{2m\omega_1}(\not\epsilon_2\not\epsilon_1\not k_1),\\ \tilde{\mathcal{O}}_{\rm a} &= \frac{1}{2m\omega_1}\gamma_0(\not\epsilon_2\not\epsilon_1\not k_1)^\dagger\gamma_0 = \frac{1}{2m\omega_1}(\not k_1\not\epsilon_1\not\epsilon_2).\end{aligned}$$

$$\begin{aligned}T_{\rm aa} &= \frac{1}{4m^2\omega_1^2}{\rm Tr}\left[\not\epsilon_2\not\epsilon_1\not k_1(\not p_1+m)\not k_1\not\epsilon_1\not\epsilon_2(\not p_2+m)\right]\\ &= \frac{1}{4m^2\omega_1^2}{\rm Tr}\left[\not\epsilon_2\not\epsilon_1\not k_1\not p_1\not k_1\not\epsilon_1\not\epsilon_2\not p_2 + m^2\not\epsilon_2\not\epsilon_1\underbrace{\not k_1\not k_1}\not\epsilon_1\not\epsilon_2\right].\end{aligned}\tag{4.206}$$

In the trace, we have omitted the cross terms involving m since they involve odd number of γ matrices and the trace of a product of an odd number of γ matrices vanishes.

The m^2 term in the trace of Eq. (4.206) vanishes since[18]

$$\not k_1\not k_1 = \mathbf{k}_1\cdot\mathbf{k}_1 = 0,\tag{4.207}$$

and we are left with only one term.

$$\begin{aligned}T_{\rm aa} &= \frac{1}{4m^2\omega_1^2}{\rm Tr}\left[\not\epsilon_2\not\epsilon_1\underbrace{\not k_1\not p_1}\not k_1\not\epsilon_1\not\epsilon_2\not p_2\right]\\ &= \frac{1}{4m^2\omega_1^2}{\rm Tr}\left[\not\epsilon_2\not\epsilon_1\{-\not p_1\not k_1 + 2(\mathbf{p}_1\cdot\mathbf{k}_1)\}\not k_1\not\epsilon_1\not\epsilon_2\not p_2\right],\end{aligned}\tag{4.208}$$

using the relation

$$\not k_1\not p_1 = -\not p_1\not k_1 + 2(\mathbf{p}_1\cdot\mathbf{k}_1).\tag{4.209}$$

Applying the relation (4.207) once again to Eq. (4.208), the first term vanishes and we are left with only the second term.

$$\begin{aligned}T_{\rm aa} &= \frac{1}{4m^2\omega_1^2}2(\mathbf{p}_1\cdot\mathbf{k}_1){\rm Tr}\left[\not\epsilon_2\underbrace{\not\epsilon_1\not k_1}\not\epsilon_1\not\epsilon_2\not p_2\right]\\ &= \frac{1}{4m^2\omega_1^2}2(\mathbf{p}_1\cdot\mathbf{k}_1){\rm Tr}\left[\not\epsilon_2\{-\not k_1\not\epsilon_1 + 2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_1)\}\not\epsilon_1\not\epsilon_2\not p_2\right]\\ &= \frac{1}{4m^2\omega_1^2}2(\mathbf{p}_1\cdot\mathbf{k}_1){\rm Tr}\left[-\not\epsilon_2\not k_1\underbrace{\not\epsilon_1\not\epsilon_1}\not\epsilon_2\not p_2\right]\\ &= \frac{1}{4m^2\omega_1^2}2(\mathbf{p}_1\cdot\mathbf{k}_1){\rm Tr}\left[\not\epsilon_2\not k_1\not\epsilon_2\not p_2\right].\end{aligned}\tag{4.210}$$

[18] We have used underbrace to highlight the relevant terms in the step-by-step evaluation.

Equation (4.210) is obtained by using the relations[19]

$$\mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_1 = 0 \quad \text{and} \quad \not{\varepsilon}_1 \not{\varepsilon}_1 = \boldsymbol{\varepsilon}_1 \cdot \boldsymbol{\varepsilon}_1 = -1. \tag{4.211}$$

We shall evaluate the trace that occurs in Eq. (4.210).

$$\begin{aligned} \mathrm{Tr}(\not{\varepsilon}_2 \not{k}_1 \not{\varepsilon}_2 \not{p}_2) &= 4(\mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_2)(\mathbf{p}_2 \cdot \boldsymbol{\varepsilon}_2) - 4(\boldsymbol{\varepsilon}_2 \cdot \boldsymbol{\varepsilon}_2)(\mathbf{p}_2 \cdot \mathbf{k}_1) \\ &\quad +4(\mathbf{p}_2 \cdot \boldsymbol{\varepsilon}_2)(\mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_2) \\ &= 8(\mathbf{p}_2 \cdot \boldsymbol{\varepsilon}_2)(\mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_2) + 4(\mathbf{p}_2 \cdot \mathbf{k}_1). \end{aligned} \tag{4.212}$$

Due to energy-momentum conservation,

$$\mathbf{p}_2 = \mathbf{p}_1 + \mathbf{k}_1 - \mathbf{k}_2. \tag{4.213}$$

Consequently,

$$\mathbf{p}_2 \cdot \boldsymbol{\varepsilon}_2 = \mathbf{p}_1 \cdot \boldsymbol{\varepsilon}_2 + \mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_2 - \mathbf{k}_2 \cdot \boldsymbol{\varepsilon}_2 = \mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_2. \tag{4.214}$$

Substituting (4.212) and (4.214) into Eq. (4.210), we get

$$T_{\mathrm{aa}} = \frac{2}{m^2 \omega_1^2} (\mathbf{p}_1 \cdot \mathbf{k}_1) \{ \mathbf{p}_2 \cdot \mathbf{k}_1 + 2(\mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_2)^2 \}. \tag{4.215}$$

Now, we shall choose the coordinate system in which the initial electron is at rest. Then

$$\mathbf{p}_1 \cdot \mathbf{k}_1 = m\omega_1 \tag{4.216}$$

$$\begin{aligned} \mathbf{p}_2 \cdot \mathbf{k}_1 &= (\mathbf{p}_1 + \mathbf{k}_1 - \mathbf{k}_2) \cdot \mathbf{k}_1 \\ &= m\omega_1 - \mathbf{k}_1 \cdot \mathbf{k}_2, \text{ since } \mathbf{k}_1 \cdot \mathbf{k}_1 = 0 \\ &= m\omega_1 - \omega_1 \omega_2 (1 - \cos\theta). \end{aligned} \tag{4.217}$$

Using the compton frequency shift formula (4.35),

$$1 - \cos\theta = \frac{m}{\omega_2} - \frac{m}{\omega_1},$$

we find

$$\mathbf{p}_2 \cdot \mathbf{k}_1 = m\omega_2. \tag{4.218}$$

Substituting (4.216) and (4.218) into Eq. (4.215), we get

$$T_{\mathrm{aa}} = \frac{2}{m\omega_1} \left\{ m\omega_2 + 2(\mathbf{k}_1 \cdot \boldsymbol{\varepsilon}_2)^2 \right\}.$$

which is the same as Eq. (4.62).

[19] $\boldsymbol{\varepsilon}_1$ has no time component and its space component is normal to the momentum vector $\boldsymbol{k}_1$.

To derive $T_{\rm bb}$

This can be done by using the same procedure outlined above. But one can straight away write down the result using the *crossing symmetry*.

$$\omega_1 \longrightarrow \omega_2; \qquad \mathbf{k}_1 \longrightarrow -\mathbf{k}_2; \qquad \boldsymbol{\varepsilon}_2 \longrightarrow \boldsymbol{\varepsilon}_1.$$

This will yield the expression (4.63) for $T_{\rm bb}$.

To derive $T_{\rm ab}$

$$T_{\rm ab} = {\rm Tr}\left[\mathcal{O}_{\rm a}(\not{p}_1 + m)\tilde{\mathcal{O}}_{\rm b}(\not{p}_2 + m)\right], \tag{4.219}$$

where

$$\begin{aligned} \mathcal{O}_{\rm a} &= \frac{1}{2m\omega_1}(\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1); \\ \tilde{\mathcal{O}}_{\rm b} &= \frac{1}{2m\omega_2}\gamma_0(\not{\varepsilon}_1\not{\varepsilon}_2\not{k}_2)^\dagger\gamma_0 = \frac{1}{2m\omega_2}(\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1). \end{aligned}$$

$$\begin{aligned} T_{\rm ab} &= \frac{1}{4m^2\omega_1\omega_2}{\rm Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1(\not{p}_1 + m)\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1(\not{p}_2 + m)\right] \\ &= \frac{1}{4m^2\omega_1\omega_2}{\rm Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\not{p}_2 + m^2\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\right] \\ &= \frac{1}{4m^2\omega_1\omega_2}[A + B], \end{aligned} \tag{4.220}$$

where

$$A = {\rm Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\not{p}_2\right]; \tag{4.221}$$

$$B = m^2{\rm Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\right]. \tag{4.222}$$

In the trace, we have omitted the cross terms involving m since they consist of odd number of γ matrices and the trace of a product of an odd number of γ matrices vanishes.

Due to energy-momentum conservation,

$$\mathbf{p}_2 = \mathbf{p}_1 + \mathbf{k}_1 - \mathbf{k}_2 \quad \text{and} \quad \not{p}_2 = \not{p}_1 + \not{k}_1 - \not{k}_2.$$

Substituting this in Eq. (4.221),

$$\begin{aligned} A &= {\rm Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\not{p}_1\right] + {\rm Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\right] \\ &\quad -{\rm Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_2\right] \\ &= A_1 + A_2 + A_3. \end{aligned} \tag{4.223}$$

First let us take A_1 and commute $\not{p}_1$ successively with the quantity on the right until it pairs with the other $\not{p}_1$. Using the result $\not{p}_1 \cdot \not{p}_1 = \mathbf{p}_1^2 = m^2$, we find that one of the terms cancels with B.

$$\begin{aligned}
A_1 &= \mathrm{Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\underbrace{\not{p}_1\not{k}_2}\not{\varepsilon}_2\not{\varepsilon}_1\not{p}_1\right] \\
&= -\mathrm{Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{k}_2\underbrace{\not{p}_1\not{\varepsilon}_2}\not{\varepsilon}_1\not{p}_1\right] + 2(\mathbf{p}_1\cdot\mathbf{k}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\underbrace{\not{\varepsilon}_1\not{k}_1}\not{\varepsilon}_2\not{\varepsilon}_1\not{p}_1\right] \\
&= \mathrm{Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{k}_2\not{\varepsilon}_2\underbrace{\not{p}_1\not{\varepsilon}_1}\not{p}_1\right] - 2(\mathbf{p}_1\cdot\mathbf{k}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\not{k}_1\underbrace{\not{\varepsilon}_1\not{\varepsilon}_2}\not{\varepsilon}_1\not{p}_1\right] \\
&= -\mathrm{Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\underbrace{\not{p}_1\not{p}_1}\right] + 2(\mathbf{p}_1\cdot\mathbf{k}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\not{k}_1\not{\varepsilon}_2\underbrace{\not{\varepsilon}_1\not{\varepsilon}_1}\not{p}_1\right] \\
&\quad -4(\mathbf{p}_1\cdot\mathbf{k}_2)(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\not{k}_1\not{\varepsilon}_1\not{p}_1\right] \\
&= -m^2\mathrm{Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\right] - 2(\mathbf{p}_1\cdot\mathbf{k}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\not{k}_1\not{\varepsilon}_2\not{p}_1\right] \\
&\quad -4(\mathbf{p}_1\cdot\mathbf{k}_2)(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\not{k}_1\not{\varepsilon}_1\not{p}_1\right] \\
&= -B - 2(\mathbf{p}_1\cdot\mathbf{k}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\not{k}_1\not{\varepsilon}_2\not{p}_1\right] - 4(\mathbf{p}_1\cdot\mathbf{k}_2)(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)\mathrm{Tr}\left[\not{\varepsilon}_2\not{k}_1\not{\varepsilon}_1\not{p}_1\right]
\end{aligned} \tag{4.224}$$

Thus

$$\begin{aligned}
A_1 + B &= -8(\mathbf{p}_1\cdot\mathbf{k}_2)\left\{(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)(\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_2) - (\boldsymbol{\varepsilon}_2\cdot\boldsymbol{\varepsilon}_2)(\mathbf{p}_1\cdot\mathbf{k}_1)\right. \\
&\quad \left. + (\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_2)(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)\right\} - 16(\mathbf{p}_1\cdot\mathbf{k}_2)(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)\left[(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)(\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_1)\right. \\
&\quad \left. -(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)(\mathbf{p}_1\cdot\mathbf{k}_1) + (\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_2)(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_1)\right] \\
&= 8(\mathbf{p}_1\cdot\mathbf{k}_2)(\boldsymbol{\varepsilon}_2\cdot\boldsymbol{\varepsilon}_2)(\mathbf{p}_1\cdot\mathbf{k}_1) + 16(\mathbf{p}_1\cdot\mathbf{k}_2)(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2(\mathbf{p}_1\cdot\mathbf{k}_1) \\
&= -8(\mathbf{p}_1\cdot\mathbf{k}_2)(\mathbf{p}_1\cdot\mathbf{k}_1) + 16(\mathbf{p}_1\cdot\mathbf{k}_2)(\mathbf{p}_1\cdot\mathbf{k}_1)(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2,
\end{aligned} \tag{4.225}$$

since $\boldsymbol{\varepsilon}_2\cdot\boldsymbol{\varepsilon}_2 = -1$.

$$\begin{aligned}
A_2 &= \mathrm{Tr}\left[\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\right] \\
&= \mathrm{Tr}\left[\not{\varepsilon}_1\underbrace{\not{k}_1\not{\varepsilon}_2}\not{\varepsilon}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\right] \\
&= -\mathrm{Tr}\left[\not{\varepsilon}_1\not{\varepsilon}_2\underbrace{\not{k}_1\not{\varepsilon}_1}\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\right] + 2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)\mathrm{Tr}\left[\underbrace{\not{\varepsilon}_1\not{\varepsilon}_1}\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\right] \\
&= \mathrm{Tr}\left[\not{\varepsilon}_1\not{\varepsilon}_2\not{\varepsilon}_1\not{k}_1\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\right] - 2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_1)\mathrm{Tr}\left[\not{\varepsilon}_1\not{\varepsilon}_2\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\right] \\
&\quad -2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)\mathrm{Tr}\left[\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\right] \\
&= -2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)\mathrm{Tr}\left[\not{k}_1\not{p}_1\not{k}_2\not{\varepsilon}_2\right] \\
&= -8(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)\left[(\mathbf{p}_1\cdot\mathbf{k}_1)(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_2) - (\mathbf{k}_1\cdot\mathbf{k}_2)(\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_2)\right. \\
&\quad \left. +(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)(\mathbf{p}_1\cdot\mathbf{k}_2)\right] \\
&= -8(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2(\mathbf{p}_1\cdot\mathbf{k}_2)
\end{aligned} \tag{4.226}$$

$$
\begin{aligned}
A_3 &= -\mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\underbrace{\not k_2\not\varepsilon_2}\not\varepsilon_1\not k_2\right] \\
&= \mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\not\varepsilon_2\not k_2\not\varepsilon_1\not k_2\right] - 2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_2)\mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\not\varepsilon_1\not k_2\right] \\
&= \mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\not\varepsilon_2\underbrace{\not k_2\not\varepsilon_1}\not k_2\right] \\
&= -\mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\not\varepsilon_2\not\varepsilon_1\underbrace{\not k_2\not k_2}\right] + 2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)\mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\not\varepsilon_2\not k_2\right] \\
&= 2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)\mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\underbrace{\not\varepsilon_2\not k_2}\right] \\
&= -2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)\mathrm{Tr}\left[\not\varepsilon_2\not\varepsilon_1\not k_1\not p_1\not k_2\not\varepsilon_2\right] \\
&= 2(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)\mathrm{Tr}\left[\not\varepsilon_1\not k_1\not p_1\not k_2\right] \\
&= 8(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)\left[(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_1)(\mathbf{p}_1\cdot\mathbf{k}_2) - (\mathbf{p}_1\cdot\boldsymbol{\varepsilon}_1)(\mathbf{k}_1\cdot\mathbf{k}_2)\right. \\
&\quad \left.+(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)(\mathbf{p}_1\cdot\mathbf{k}_1)\right] \\
&= 8(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2(\mathbf{p}_1\cdot\mathbf{k}_1). \qquad (4.227)
\end{aligned}
$$

Substituting (4.224) - (4.227) into Eq. (4.220), we get

$$
\begin{aligned}
T_{\mathrm{ab}} &= \frac{2}{m^2\omega_1\omega_2}\left[(\mathbf{p}_1\cdot\mathbf{k}_2)(\mathbf{p}_1\cdot\mathbf{k}_1)\{2(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2-1\} - (\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2(\mathbf{p}_1\cdot\mathbf{k}_2)\right. \\
&\quad \left.+(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2(\mathbf{p}_1\cdot\mathbf{k}_1)\right] \\
&= \frac{2}{m^2\omega_1\omega_2}\left[m^2\omega_1\omega_2\{2(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2-1\} - m\omega_2(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2\right. \\
&\quad \left.+m\omega_1(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2\right] \\
&= 2\{2(\boldsymbol{\varepsilon}_1\cdot\boldsymbol{\varepsilon}_2)^2-1\} - \frac{2}{m\omega_1}(\mathbf{k}_1\cdot\boldsymbol{\varepsilon}_2)^2 + \frac{2}{m\omega_2}(\mathbf{k}_2\cdot\boldsymbol{\varepsilon}_1)^2.
\end{aligned}
$$

This is the same expression as given by Eq. (4.64)

4.3 Let us start with the Klein-Nishina formula.

$$
\frac{d\sigma}{d\Omega} = \frac{e^4}{2m^2}\left[\left(\frac{\omega_2}{\omega_1}\right)^3 + \frac{\omega_2}{\omega_1} - \left(\frac{\omega_2}{\omega_1}\right)^2(1-\cos^2\theta)\right].
$$

Since ω_2 depends on θ as shown in Eqs. (4.74) and (4.75), we need to take into account its dependence on θ while integrating. We shall perform the integration of the terms one by one.

First term

$$
\int\left(\frac{\omega_2}{\omega_1}\right)^3 d\Omega = 2\pi\int_0^\pi \frac{\sin\theta d\theta}{(1+r-r\cos\theta)^3} = 2\pi\int_{-1}^{+1}\frac{dx}{(a+bx)^3},
$$

with

$$
a = 1+r, \qquad b = -r, \qquad x = \cos\theta.
$$

Using the standard integral,

$$\int \frac{dx}{(a+bx)^3} = -\frac{1}{2b(a+bx)^2},$$

and substituting the limits, we obtain

$$\int \left(\frac{\omega_2}{\omega_1}\right)^3 d\Omega = 4\pi \frac{1+r}{(1+2r)^2}. \tag{4.228}$$

Second term

$$\int \left(\frac{\omega_2}{\omega_1}\right) d\Omega = 2\pi \int_0^\pi \frac{\sin\theta d\theta}{1+r-r\cos\theta} = 2\pi \int_{-1}^{+1} \frac{dx}{a+bx},$$

Using the standard integral,

$$\int \frac{dx}{a+bx} = \frac{1}{b}\ln(a+bx),$$

and substituting the limits, we obtain

$$\int \left(\frac{\omega_2}{\omega_1}\right) d\Omega = \frac{2\pi}{r}\ln(1+2r). \tag{4.229}$$

Third term

$$\int \left(\frac{\omega_2}{\omega_1}\right)^2 d\Omega = 2\pi \int_0^\pi \frac{\sin\theta d\theta}{(1+r-r\cos\theta)^2} = 2\pi \int_{-1}^{+1} \frac{dx}{(a+bx)^2},$$

Using the standard integral,

$$\int \frac{dx}{(a+bx)^2} = -\frac{1}{b(a+bx)},$$

and substituting the limits, we obtain

$$\int \left(\frac{\omega_2}{\omega_1}\right)^2 d\Omega = \frac{4\pi}{1+2r}. \tag{4.230}$$

Fourth term

$$\int \left(\frac{\omega_2}{\omega_1}\right)^2 \cos^2\theta d\Omega = 2\pi \int_0^\pi \frac{\cos^2\theta\sin\theta d\theta}{(1+r-r\cos\theta)^2} = 2\pi \int_{-1}^{+1} \frac{x^2 dx}{(a+bx)^2},$$

Using the standard integral,

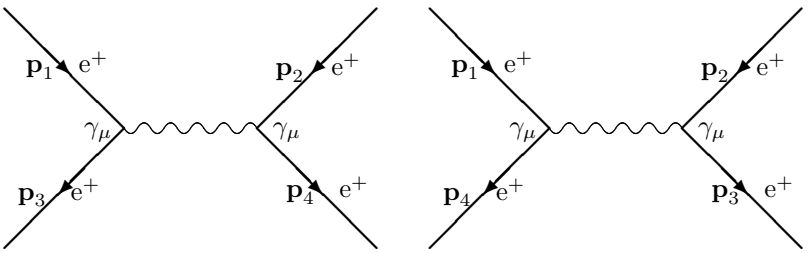
$$\int \frac{x^2 dx}{(a+bx)^2} = \frac{1}{b^3}\left\{a+bx-2a\ln(a+bx)-\frac{a^2}{a+bx}\right\},$$

and substituting the limits, we obtain

$$\int \left(\frac{\omega_2}{\omega_1}\right)^2 \cos^2\theta d\Omega = 4\pi \left\{\frac{1}{r^2} - \frac{1+r}{r^3}\ln(1+2r) + \frac{(1+r)^2}{r^2(1+2r)}\right\}. \quad (4.231)$$

Collecting all these results, we get the total cross section given by Eq. (4.76).

4.4 Feynman diagrams representing the positron-positron scattering is given below. Due to indistinguishability of positrons in the final state, one has to consider both diagrams (a) and (b).



(a) (b)

Figure 4.12: Positron-positron scattering.

The matrix element for the process can easily be written down in the momentum representation by using the Feynman rules. One has to follow the direction of arrows in writing down the matrix elements. For the sake of brevity, we denote the spinors $v(\mathbf{p}_1), v(\mathbf{p}_2)\cdots$ by $v_1, v_2, \cdots$.

$$\mathcal{M} = i4\pi e^2 \left[\frac{(\bar{v}_4\gamma_\mu v_2)(\bar{v}_3\gamma_\mu v_1)}{(\mathbf{p}_1 - \mathbf{p}_3)^2} - \frac{(\bar{v}_3\gamma_\mu v_2)(\bar{v}_4\gamma_\mu v_1)}{(\mathbf{p}_1 - \mathbf{p}_4)^2}\right], \quad (4.232)$$

where summation over μ is implied. The first term in Eq. (4.232) corresponds to the diagram (a) and the second term corresponds to the exchange diagram (b). The negative sign between the two terms arises from the antisymmetry of the two-positron wave function. To find the transition probability, we have to find the absolute square of the matrix element (4.232). If the positron spins are not observed, then a sum over the final spin states and an average over the initial spin states have to be taken.

This introduces a factor 1/4.

$$
\begin{aligned}
|\overline{\mathcal{M}}|^2 &= \frac{1}{4}\mathcal{M}\mathcal{M}^\dagger \\
&= 4\pi^2 e^4 \sum_{\text{spins}} \left[\frac{(\bar{v}_4\gamma_\mu v_2)(\bar{v}_3\gamma_\mu v_1)}{(\mathbf{p}_1-\mathbf{p}_3)^2} - \frac{(\bar{v}_3\gamma_\mu v_2)(\bar{v}_4\gamma_\mu v_1)}{(\mathbf{p}_1-\mathbf{p}_4)^2}\right] \\
&\quad \times \left[\frac{(\bar{v}_4\gamma_\nu v_2)(\bar{v}_3\gamma_\nu v_1)}{(\mathbf{p}_1-\mathbf{p}_3)^2} - \frac{(\bar{v}_3\gamma_\nu v_2)(\bar{v}_4\gamma_\nu v_1)}{(\mathbf{p}_1-\mathbf{p}_4)^2}\right]^\dagger \\
&= 4\pi^2 e^4 \left[\frac{T_{\text{aa}}}{(\mathbf{p}_1-\mathbf{p}_3)^4} + \frac{T_{\text{bb}}}{(\mathbf{p}_1-\mathbf{p}_4)^4} - \frac{T_{\text{ab}}+T_{\text{ba}}}{(\mathbf{p}_1-\mathbf{p}_3)^2(\mathbf{p}_1-\mathbf{p}_4)^2}\right],
\end{aligned}
$$

where

$$
\begin{aligned}
T_{\text{aa}} &= \sum_{\text{spins}} \left[(\bar{v}_3\gamma_\mu v_1)(\bar{v}_3\gamma_\nu v_1)^\dagger(\bar{v}_4\gamma_\mu v_2)(\bar{v}_4\gamma_\nu v_2)^\dagger\right] \\
&= \text{Tr}\left[\gamma_\mu(\not{p}_1 - m)\gamma_\nu(\not{p}_3 - m)\right]\text{Tr}\left[\gamma_\mu(\not{p}_2 - m)\gamma_\nu(\not{p}_4 - m)\right], \\
T_{\text{bb}} &= \sum_{\text{spins}} \left[(\bar{v}_4\gamma_\mu v_1)(\bar{v}_4\gamma_\nu v_1)^\dagger(\bar{v}_3\gamma_\mu v_2)(\bar{v}_3\gamma_\nu v_2)^\dagger\right] \\
&= \text{Tr}\left[\gamma_\mu(\not{p}_1 - m)\gamma_\nu(\not{p}_4 - m)\right]\text{Tr}\left[\gamma_\mu(\not{p}_2 - m)\gamma_\nu(\not{p}_3 - m)\right], \\
T_{\text{ab}} &= \sum_{\text{spins}} \left[(\bar{v}_3\gamma_\mu v_1)(\bar{v}_4\gamma_\nu v_1)^\dagger(\bar{v}_4\gamma_\mu v_2)(\bar{v}_3\gamma_\nu v_2)^\dagger\right] \\
&= \text{Tr}\left[\gamma_\mu(\not{p}_1 - m)\gamma_\nu(\not{p}_4 - m)\gamma_\mu(\not{p}_2 - m)\gamma_\nu(\not{p}_3 - m)\right], \\
T_{\text{ba}} &= \sum_{\text{spins}} \left[(\bar{v}_4\gamma_\mu v_1)(\bar{v}_3\gamma_\nu v_1)^\dagger(\bar{v}_3\gamma_\mu v_2)(\bar{v}_4\gamma_\nu v_2)^\dagger\right] \\
&= \text{Tr}\left[\gamma_\mu(\not{p}_1 - m)\gamma_\nu(\not{p}_3 - m)\gamma_\mu(\not{p}_2 - m)\gamma_\nu(\not{p}_4 - m)\right].
\end{aligned}
$$

The expressions for $T_{\text{aa}}, T_{\text{bb}}, T_{\text{ab}}, T_{\text{ba}}$ given above have been obtained using the projection operators for positrons

$$\sum_{\text{spins}} v\bar{v} = \not{p} - m$$

and they are identical with the expressions (4.83) - (4.86) obtained earlier for electron-electron scattering except for the change in sign for m. Since the traces of a product of odd number of gamma matrices vanish, one has to consider the terms involving even number of gamma matrices only. That means one has to consider only products of even number of m and so the results will be independent of the sign of m.

Hence, it follows that the positron-positron scattering cross section will be identical with electron-electron scattering cross section.

Chapter 5

Radiative Corrections

In the last chapter, we have studied the various processes in QED in the lowest order of perturbation theory. The higher order terms are expected to yield small corrections of the order of the fine structure constant α. But on doing such a calculation, one encounters divergent integrals.

In the lowest order, the momenta of the particles in the intermediate stage are well defined because of the conservation of momenta at every vertex. But in higher order, the momenta of the intermediate particles can vary over a wide range and this causes divergences. Special techniques are to be devised to overcome these difficulties. These techniques have come to be known as *Regularization* and *Renormalization*. They are based on the arguments that what one considers in perturbation theory is the bare electron with mass m_0 and charge $-e_0$ and what one observes in experiments is the physical electron with mass m and charge $-e$.

The higher order corrections are sometimes called *Radiative corrections* and they can be broadly classified into three types, electron self energy, photon self energy or vacuum polarization and vertex corrections. Any higher order correction to a given process will involve all these types and so it is sufficient to study them separately and devise methods to overcome the divergences.

5.1 Electron Self Energy

The Feynman diagram illustrating the mechanism by which an electron acquires self-energy is given in Fig. 5.1. The incident electron of mo-

mentum $\boldsymbol{p}$ emits a virtual photon of momentum $\boldsymbol{k}$ which is recaptured by the emerging electron of the same momentum $\boldsymbol{p}$. The virtual photon momentum $\boldsymbol{k}$ can assume any value from 0 to ∞ and hence an integration has to be performed over the momenta and this results in the divergence problem.

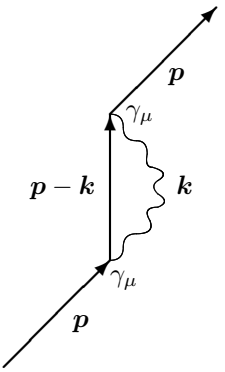


Figure 5.1: Self energy of the electron

The matrix element[1] for this process in momentum representation can be written down[2] using the Feynman rules.

$$\mathcal{M} = -i4\pi e^2 \int \bar{u}(\boldsymbol{p})\gamma_\mu \frac{1}{\not{p} - \not{k} - m}\gamma_\mu u(\boldsymbol{p}) \frac{d^4k}{(2\pi)^4}\frac{1}{\mathbf{k}^2}. \tag{5.1}$$

Since the initial and final momenta of the electron are identical, the matrix element can be interpreted as the change in energy of the electron due to self-interaction. Introducing the normalization factor $\frac{1}{2E}$ for the electron spinors, we obtain the following expression for the change in energy due to self interaction.

$$\Delta E = -i\frac{4\pi e^2}{2E} \int \bar{u}(\boldsymbol{p})\gamma_\mu \frac{1}{\not{p} - \not{k} - m}\gamma_\mu u(\boldsymbol{p}) \frac{d^4k}{(2\pi)^4}\frac{1}{\mathbf{k}^2}. \tag{5.2}$$

Since the momentum of the electron remains the same after the perturbation, the change in energy can be interpreted as the change in mass of the electron.

$$E^2 = p^2 + m^2; \qquad 2EdE = 2mdm.$$

[1] R. P. Feynman, *Quantum Electrodynamics*, W. A. Benjamin, Inc. (1962).

[2] We use upright bold letters $(\mathbf{k}, \mathbf{p})$ to denote four-vectors, italic bold letters $(\boldsymbol{k}, \boldsymbol{p})$ to denote three-vectors. The symbols $\not{p}$, $\not{k}$ represent $\not{p} = \gamma_\mu p_\mu$, $\not{k} = \gamma_\mu k_\mu$.

Thus, we obtain an expression for the change in mass of the electron due to self-interaction.

$$\Delta m = \frac{4\pi e^2}{2mi}\int \bar{u}\left(\gamma_\mu \frac{1}{\not p - \not k - m}\gamma_\mu\right) u \frac{d^4k}{(2\pi)^4}\frac{1}{\mathbf{k}^2}. \tag{5.3}$$

The quantity within the bracket can be simplified to yield

$$\gamma_\mu \frac{1}{\not p - \not k - m}\gamma_\mu = \frac{\gamma_\mu(\not p - \not k + m)\gamma_\mu}{(\not p - \not k)^2 - m^2} = \frac{2\not k + 2m}{\mathbf{k}^2 - 2\mathbf{p}\cdot\mathbf{k}}, \tag{5.4}$$

using the following relations.

$$\begin{aligned} \gamma_\mu \not p \gamma_\mu &= -2\not p \to -2m, \qquad \text{since } \not p u = mu; \\ \gamma_\mu \not k \gamma_\mu &= -2\not k; \qquad \gamma_\mu\gamma_\mu = 4; \qquad \not p^2 = \mathbf{p}^2 = m^2. \end{aligned}$$

Equation (5.3) now becomes

$$\Delta m = \frac{4\pi e^2}{2mi}\int \frac{\bar{u}(2m + 2\not k)u}{\mathbf{k}^2 - 2\mathbf{p}\cdot\mathbf{k}}\frac{d^4k}{(2\pi)^4}\frac{1}{\mathbf{k}^2}. \tag{5.5}$$

This integral is divergent and it has posed a major challenge in the study of QED for over 20 years. We shall closely follow Feynman's method of overcoming this problem. Let us modify the propagation kernel of the photon by multiplying it by a factor $c(\mathbf{k}^2)$.

$$\frac{1}{\mathbf{k}^2} \to \frac{1}{\mathbf{k}^2}c(\mathbf{k}^2) = \frac{1}{\mathbf{k}^2}\left(\frac{-\lambda^2}{\mathbf{k}^2 - \lambda^2}\right) = \frac{1}{\mathbf{k}^2} - \frac{1}{\mathbf{k}^2 - \lambda^2}. \tag{5.6}$$

The second term corresponds to the propagation of photon of mass λ.

Introducing this kernel into Eq. (5.5), we get

$$\Delta m = \frac{4\pi e^2}{2mi}\int \frac{\bar{u}(2m + 2\not k)u}{\mathbf{k}^2 - 2\mathbf{p}\cdot\mathbf{k}}\frac{d^4k}{(2\pi)^4}\frac{1}{\mathbf{k}^2}\left(\frac{-\lambda^2}{\mathbf{k}^2 - \lambda^2}\right)\,. \tag{5.7}$$

A more convenient representation for the modified kernel is

$$-\int_0^{\lambda^2}\frac{dL}{(\mathbf{k}^2 - L)^2} = -\frac{1}{\mathbf{k}^2 - L}\bigg|_{L=0}^{L=\lambda^2} = -\frac{1}{\mathbf{k}^2 - \lambda^2} + \frac{1}{\mathbf{k}^2}. \tag{5.8}$$

Using (5.8), we can rewrite Eq. (5.7) in the following form.

$$\Delta m = -\frac{4\pi e^2}{2mi}\int_0^{\lambda^2} dL \int \frac{\bar{u}(2m + 2\not k)u}{(\mathbf{k}^2 - 2\mathbf{p}\cdot\mathbf{k})(\mathbf{k}^2 - L)^2}\frac{d^4k}{(2\pi)^4}. \tag{5.9}$$

The denominator in Eq. (5.9) can be expressed as

$$\frac{1}{(\mathbf{k}^2 - 2\mathbf{p}\cdot\mathbf{k})(\mathbf{k}^2 - L)^2} = \int_0^1 \frac{2(1-x)dx}{[\mathbf{k}^2 - 2x\mathbf{p}\cdot\mathbf{k} - L(1-x)]^3}, \tag{5.10}$$

using the remarkable Feynman's formula (*vide* Solved Problem 5.2)

$$\frac{1}{ab^2} = \int_0^1 \frac{2(1-x)dx}{[ax + b(1-x)]^2}, \tag{5.11}$$

with the substitution

$$a = \mathbf{k}^2 - 2\mathbf{p}\cdot\mathbf{k}, \qquad b = \mathbf{k}^2 - L.$$

Substituting (5.10) into Eq. (5.9), we get

$$\Delta m = -\frac{4\pi e^2}{2mi}\int_0^{\lambda^2} dL \int_0^1 dx \int \frac{d^4k}{(2\pi)^4} \frac{2(1-x)\bar{u}(2m + 2\not{k})u}{[\mathbf{k}^2 - 2x\mathbf{p}\cdot\mathbf{k} - L(1-x)]^3}. \tag{5.12}$$

There are, in total, three integrations and let us perform them one by one.

Integration over d^4k

The d^4k integration can be performed using the following formula (*vide* Solved Problems (5.3) and (5.4)),

$$\int \frac{(1, k_\sigma)}{(\mathbf{k}^2 - 2\mathbf{p}\cdot\mathbf{k} - \Delta)^3} \frac{d^4k}{(2\pi)^4} = \frac{(1, p_\sigma)}{32\pi^2 i(\mathbf{p}^2 + \Delta)}, \tag{5.13}$$

which is given in a compact form. The notation $(1, k_\sigma)$ on the L.H.S. means either 1 or k_σ and it correspondingly yields either 1 or p_σ on the R.H.S. represented by the notation $(1, p_\sigma)$.

By making the substitution

$$\mathbf{p} \to x\mathbf{p} \quad \text{and} \quad \Delta \to L(1-x),$$

we get

$$\begin{aligned} \int \frac{(1, k_\sigma)}{[\mathbf{k}^2 - 2x\mathbf{p}\cdot\mathbf{k} - L(1-x)]^3} \frac{d^4k}{(2\pi)^4} &= \frac{(1, xp_\sigma)}{32\pi^2 i[x^2\mathbf{p}^2 + L(1-x)]} \\ &= \frac{(1, xp_\sigma)}{32\pi^2 i[m^2x^2 + L(1-x)]}, \end{aligned} \tag{5.14}$$

since $\mathbf{p}^2 = m^2$.

Integration over dL

Taking note of the factors that depend upon L, we get

$$\begin{aligned}\int_0^{\lambda^2} \frac{dL}{m^2x^2 + L(1-x)} &= \frac{\ln[m^2x^2 + L(1-x)]}{1-x}\Bigg|_{L=0}^{L=\lambda^2} \\ &= \frac{1}{1-x}\ln\left[\frac{m^2x^2 + \lambda^2(1-x)}{m^2x^2}\right]. \quad (5.15)\end{aligned}$$

When $\lambda^2 \gg m^2$, m^2x^2 occurring in the numerator can be neglected. Of course, when $x \approx .1$, $(1-x)\lambda^2$ is not much larger than m^2x^2, but the interval over which this is true is so small, for $\lambda^2 \gg m^2$, that the error is small when the x-integration is performed. Under this approximation,

$$\int_0^{\lambda^2} \frac{dL}{m^2x^2 + L(1-x)} \approx \frac{1}{1-x}\ln\left[\frac{\lambda^2(1-x)}{m^2x^2}\right]. \quad (5.16)$$

Substituting (5.14) and (5.16) into Eq. (5.12), we get

$$\begin{aligned}\Delta m &= -\frac{4\pi e^2}{2mi}\frac{1}{16\pi^2 i}\int_0^1 dx\{\bar{u}(2m + 2x\not{p})u\}\ln\left(\frac{\lambda^2(1-x)}{m^2x^2}\right) \\ &= \frac{e^2}{8\pi m}\int_0^1 dx\{\bar{u}(2m + 2x\not{p})u\ln\left(\frac{\lambda^2(1-x)}{m^2x^2}\right) \\ &= \frac{e^2m}{2\pi}\int_0^1 (1+x)\ln\left(\frac{\lambda^2(1-x)}{m^2x^2}\right). \quad (5.17)\end{aligned}$$

To obtain the last step, we have used the relations $\not{p}u = mu$ and $\bar{u}u = 2m$.

Integration over dx

Expanding the ln term

$$\ln\left(\frac{\lambda^2(1-x)}{m^2x^2}\right) = \ln\frac{\lambda^2}{m^2} + \ln\frac{1-x}{x^2} = 2\ln\frac{\lambda}{m} + \ln\frac{1-x}{x^2},$$

and using the definite integrals (*vide* Solved Problem 5.5)

$$\int_0^1 \ln\left(\frac{1-x}{x^2}\right) = 1; \qquad \int_0^1 x\ln\left(\frac{1-x}{x^2}\right) = -1/4; \quad (5.18)$$

we get a simple expression for the change in mass due to self interaction.

$$\frac{\Delta m}{m} = \frac{e^2}{2\pi}\left[3\ln\frac{\lambda}{m} + \frac{3}{4}\right]. \quad (5.19)$$

The factor $(e^2/2\pi)$ is of the order of 10^{-3}. Even if λ is many times m, the fractional change in mass will not be very large. However, the shift in mass depends upon the parameter λ and hence cannot be determined theoretically. The theoretical and the experimental mass[3] are related by

$$m_{\text{exp}} = m_{\text{th}} + \delta m.$$

All our experimental measurements are made with experimental mass with self-action included and the theoretical mass (mass without self-action) cannot be determined. The replacement of theoretical mass m_{th} by the experimental mass m_{exp} is what is known as mass renormalization and m_{exp} is often referred to as the renormalized mass.

For a free particle, a theory using m_{th} plus self-action is equivalent to a theory using m_{exp} + self-action minus Δm. For a bound-state electron, a theory using m_{exp} will slightly differ from this principle of equivalence and this leads to the Lamb shift in the hydrogen atom.

5.2 Higher Order Corrections

For the sake of illustration, let us consider the scattering of an electron by a static electric potential such as the Coulomb scattering by a nucleus of charge ze. Figure 5.2 represents the lowest order Feynman diagram.

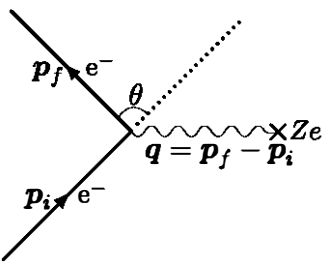


Figure 5.2: Lowest order Feynman diagram representing the scattering of an electron by a static external potential (for example, Coulomb potential by a nucleus of charge Ze). $\boldsymbol{p}_i$ and $\boldsymbol{p}_f$ denote the momenta of the incoming and scattered electron and $\boldsymbol{q}$, the three-momentum transfer to the electron.

The Feynman amplitude corresponding to the lowest order diagram 5.2 is given by

$$\mathcal{M} = -ie\bar{u}(\boldsymbol{p}_f)\not{A}(\mathbf{q})u(\boldsymbol{p}_i), \qquad \boldsymbol{q} = \boldsymbol{p}_f - \boldsymbol{p}_i. \tag{5.20}$$

[3]The theoretical mass m_{th} is the bare mass m_0 without self-interaction and the experimental mass m_{exp} is the observed mass m of the electron.

The next higher order Feynman diagram for this process involves two additional vertices representing the emission and reabsorption of virtual photon as shown in fig. 5.3.

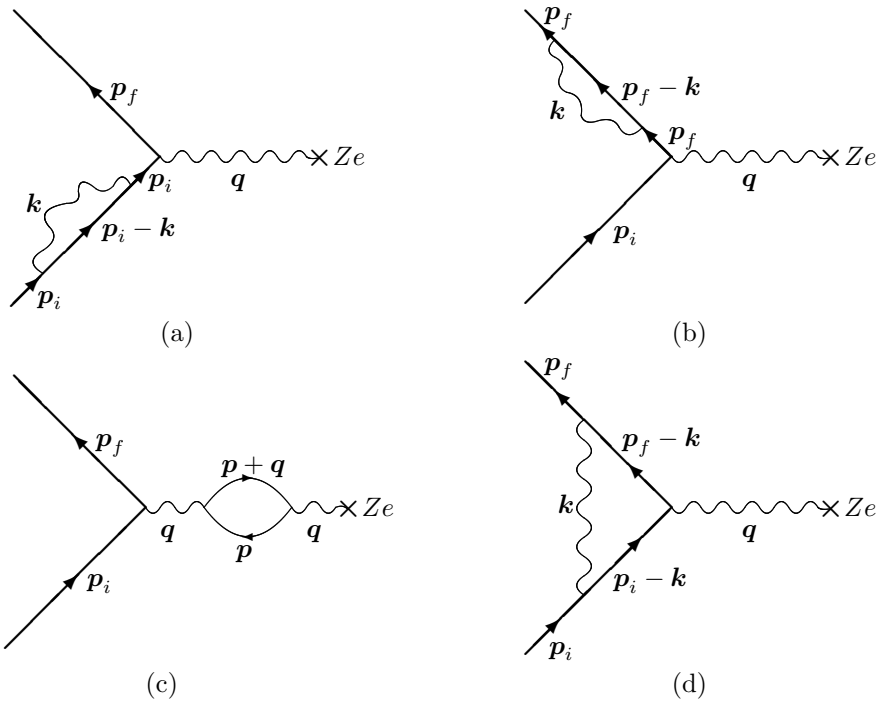


Figure 5.3: Radiative corrections to the scattering of an electron by a static external potential. $\boldsymbol{p}_i$ and $\boldsymbol{p}_f$ denote the momenta of the incoming and scattered electron and $\boldsymbol{q}$, the three-momentum transfer to the electron.

Restricting oneself to the Feynman diagrams which contain two extra vertices, involving only one virtual photon, we obtain the radiative corrections to the process of scattering of electron by the static electric field. As shown in Fig. 5.3, there are four possibilities that correspond to second order corrections to the basic process represented by Fig. 5.1. Each of them can be considered as a modification of the lowest order diagram shown in Fig. 5.1 by the substitution of one of the three loop diagrams depicted in Fig. 5.4. Diagram 5.4 (a) is the electron self-energy diagram, Diagram 5.4 (b) corresponds to the photon self-energy which is commonly referred to as the vacuum polarization diagram and Diagram 5.4 (c) corresponds to the modification of the basic vertex diagram representing Fermion-photon interaction. Each of the diagrams in Fig. 5.3

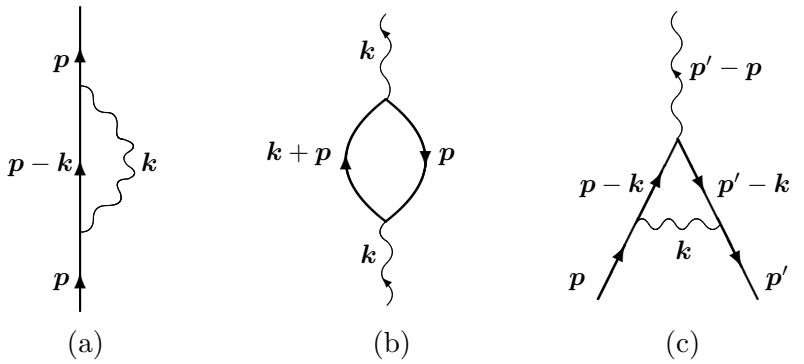


Figure 5.4: Feynman diagrams representing (a) Self energy of the electron (b) Self energy of the photon or Vacuum polarization and (c) Vertex correction.

uses one of the basic modifications depicted in Fig. 5.4.

Diagrams 5.3(a) and 5.3(b) include the basic self-energy diagram 5.4(a); Diagram 5.4(c) includes the basic photon self-energy diagram 5.4(b) and Diagram 5.4(d) includes the basic modification in the electron-photon vertex. It is possible to extract the effect of these loop diagrams depicted in Fig. 5.4 on higher order terms in perturbation theory. The loop diagrams lead to divergent integrals and it is a major problem to regularize these integrals and extract meaningful results from them. We have already seen how one can obtain meaningful results from the self-energy diagram.

$$\mathcal{M}_3^{(a)} = -4\pi e^3 \int \left(\bar{u}_f \not{\epsilon} \frac{1}{\not{p}_i - m} \gamma_\mu \frac{1}{\not{p}_i - \not{k} - m} \gamma_\mu u_i \right) \frac{1}{\mathbf{k}}^2 \frac{d^4k}{(2\pi)^4}. \quad (5.21)$$

$$\mathcal{M}_3^{(b)} = -4\pi e^3 \int \left(\bar{u}_f \gamma_\mu \frac{1}{\not{p}_f - \not{k} - m} \gamma_\mu \frac{1}{\not{p}_f - m} \not{\epsilon} u_i \right) \frac{1}{\mathbf{k}}^2 \frac{d^4k}{(2\pi)^4}. \quad (5.22)$$

$$\mathcal{M}_3^{(c)} = -4\pi e^3 (\bar{u}_f \gamma_\mu u_1) \frac{1}{\mathbf{q}^2} \int \sum_{\text{spins}} \left(\bar{u} \frac{1}{\not{p} - m} \gamma_\mu \frac{1}{\not{p} + \not{q} - m} \not{\epsilon} u \right) \frac{d^4p}{(2\pi)^4}; \quad (5.23)$$

$$\mathcal{M}_3^{(d)} = -4\pi e^3 \int \left(\bar{u}_f \gamma_\mu \frac{1}{\not{p}_f - \not{k} - m} \not{\epsilon} \frac{1}{\not{p}_i - \not{k} - m} \gamma_\mu u_i \right) \frac{1}{\mathbf{k}}^2 \frac{d^4k}{(2\pi)^4}. \quad (5.24)$$

Comparing the higher order matrix elements (5.21) - (5.24) with the lowest order matrix element (5.20), we can rewrite Eqs. (5.21) - (5.24) as given

below:

$$\mathcal{M}_3^{(a)} = -ie\bar{u}_f \not{A}\Sigma(\boldsymbol{p}_i)u_i, \tag{5.25}$$

$$\mathcal{M}_3^{(b)} = -ie\bar{u}_f \Sigma(\boldsymbol{p}_f)\not{A}u_i, \tag{5.26}$$

$$\mathcal{M}_3^{(c)} = -ie\bar{u}_f \gamma_\mu u_i \Pi, \tag{5.27}$$

$$\mathcal{M}_3^{(d)} = -ie\bar{u}_f \gamma_\mu \Lambda u_i, \tag{5.28}$$

where

$$\Sigma(\boldsymbol{p}_i) = -i4\pi e^2 \int \left(\frac{1}{\not{p}_i - m}\gamma_\mu \frac{1}{\not{p}_i - \not{k} - m}\gamma_\mu u_i\right) \frac{1}{\mathbf{k}}^2 \frac{d^4k}{(2\pi)^4}; \tag{5.29}$$

$$\Sigma(\boldsymbol{p}_f) = -i4\pi e^2 \int \left(\frac{1}{\not{p}_f - \not{k} - m}\gamma_\mu \frac{1}{\not{p}_f - m}\gamma_\mu u_i\right) \frac{1}{\mathbf{k}}^2 \frac{d^4k}{(2\pi)^4}; \tag{5.30}$$

$$\Pi(\boldsymbol{q}) = -i4\pi e^2 \int \sum_{\text{spins}} \frac{1}{\mathbf{q}^2} \left(\bar{u}\frac{1}{\not{p} - m}\gamma_\mu \frac{1}{\not{p} + \not{A} - m}\not{A}u\right) \frac{d^4p}{(2\pi)^4}; \tag{5.31}$$

$$\Lambda(\boldsymbol{p}_f, \boldsymbol{p}_i) = -i4\pi e^2 \int \left(\gamma_\mu \frac{1}{\not{p}_f - \not{k} - m}\not{A}\frac{1}{\not{p}_i - \not{k} - m}\gamma_\mu\right) \frac{1}{\mathbf{k}}^2 \frac{d^4k}{(2\pi)^4}. \tag{5.32}$$

Thus, for calculating the second order corrections to the elastic scattering of electron by a static external potential, we need to evaluate only the three loop integrals Σ, Π, Λ that occurs in expressions (5.29) - (5.31) and depicted in Fig. 5.4. All the three integrals are found to be divergent for large values of the momentum variables of integration. The concepts of mass and charge renormalization enable one to extract unambiguously finite radiative corrections of order α, expressed in terms of observed mass and charge. The importance of this analysis is to point out that the radiative corrections of lowest order to any process involve the same three divergent integrals and once we have coped up with these three integrals, the calculations of the second order radiative corrections to any process presents no difficulties, in principle.

Review Questions

5.1 Draw a Feynman diagram for calculating the self energy of electron and show that it involves divergent integrals. Explain the Feynman method of coping with this integral and show that the electron self-energy is equivalent to a shift in mass. Obtain an expression for the mass shift.

5.2 Explain the concept of mass renormalizaion and explain how one can obtain the Lamb shift for a bound electron in hydrogen atom.

5.3 Consider the elastic scattering of an electron by a static external potential and draw Feynman diagrams for obtaining second-order radiative corrections for this process.

Problems

5.1 Show that for two operators A and B, whether commuting or non-commuting, the following expansions are true:

$$\begin{aligned}\frac{1}{A-B} &= \frac{1}{A}+\frac{1}{A}B\frac{1}{A}+\frac{1}{A}B\frac{1}{A}B\frac{1}{A}+\cdots.\\ \frac{1}{A+B} &= \frac{1}{A}-\frac{1}{A}B\frac{1}{A}+\frac{1}{A}B\frac{1}{A}B\frac{1}{A}-\cdots.\end{aligned}$$

5.2 Prove the following Feynman's formulae:

$$\begin{aligned}(1)\quad \frac{1}{ab} &= \int_0^1 \frac{dz}{[az+b(1-z)]^2};\\ (2)\quad \frac{1}{a^2b} &= \int_0^1 \frac{2zdz}{[az+b(1-z)]^3};\\ (3)\quad \frac{1}{ab^2} &= \int_0^1 \frac{2(1-z)dz}{[az+b(1-z)]^3}.\end{aligned}$$

5.3 Establish the following relation

$$\int_{-\infty}^{\infty}\frac{(1,k_\sigma)d^4k}{(2\pi)^4(\mathbf{k}^2+i\epsilon-L)^3}=\frac{(1,0)}{32\pi^2iL},$$

which is given in a compact form. The notation $(1:k_\sigma)$ in the numerator means either 1 or k_σ and correspondingly the notation (1:0) on the right hand side of the equation means 1 or 0.

5.4 Given the following relation

$$\int_{-\infty}^{\infty}\frac{(1,k_\sigma)d^4k}{(2\pi)^4(\mathbf{k}^2+i\epsilon-L)^3}=\frac{(1,0)}{32\pi^2iL},$$

change the variable $\mathbf{k}\rightarrow\mathbf{k}-\mathbf{p}$ and obtain the relation given below.

$$\int_{-\infty}^{\infty}\frac{(1;k_\sigma)d^4k}{(2\pi)^4(\mathbf{k}^2-2\mathbf{p}\cdot\mathbf{k}-\Delta)^3}=\frac{(1;p_\sigma)}{32\pi^2i(\mathbf{p}^2+\Delta)},$$

5.5 Evaluate the following definite integrals and show that they yield the results given below:

$$\text{(a)}\ \int_0^1 \ln[\mathrm{x}^{-2}(1-\mathrm{x})]\mathrm{dx}=1;\qquad \text{(b)}\ \int_0^1 \mathrm{x}\,\ln[\mathrm{x}^{-2}(1-\mathrm{x})]\mathrm{dx}=-1/4.$$

Solutions to Problems

5.1 Post-multiply the first of equations by $A - B$.

$$\begin{aligned}\frac{1}{A-B}(A-B) &= \frac{1}{A}(A-B)+\frac{1}{A}B\frac{1}{A}(A-B)+\frac{1}{A}B\frac{1}{A}B\frac{1}{A}(A-B)\\ &\quad +\cdots .\\ 1 &= C_1+C_2+C_3+\cdots , \qquad (5.33)\end{aligned}$$

where

$$\begin{aligned}C_1 &= 1-\frac{1}{A}B\\ C_2 &= \frac{1}{A}B\frac{1}{A}(A-B)\\ &= \frac{1}{A}B\left(1-\frac{1}{A}B\right)\\ &= \frac{1}{A}B-\frac{1}{A}B\frac{1}{A}B\\ C_3 &= \frac{1}{A}B\frac{1}{A}B\frac{1}{A}(A-B)\\ &= \frac{1}{A}B\frac{1}{A}B-\frac{1}{A}B\frac{1}{A}B\frac{1}{A}B\end{aligned}$$

On the right-hand side of Eq. (5.33), only the first term (unity) remains and all the succeeding terms mutually cancel away.

Thus we have proved the first relation. Just change the sign of B in the relation 1 and we get the second relation.

5.2 Let us start with a simple identity.

$$\frac{1}{ab}=\frac{1}{b-a}\left(\frac{1}{a}-\frac{1}{b}\right)=\frac{1}{b-a}\int_a^b\frac{dx}{x^2}. \qquad (5.34)$$

Let $x = az + b(1 - z)$, then $dx = (a - b)dz$. Substituting them on the right hand side of Eq. (5.34), we get

$$\frac{1}{ab}=\frac{1}{b-a}\int\frac{(a-b)dz}{[az+b(1-z)]^2} \qquad (5.35)$$

The lower limit of integration $x = a$ corresponds to $z = 1$ and the upper limit $x = b$ corresponds to $z = 0$. Hence we get the relation (1)

$$\frac{1}{ab} = \int_0^1\frac{dz}{[az+b(1-z)]^2}.$$

Differentiating both sides of the above equation with respect to a, we get the relation (2).

$$\frac{1}{a^2b} = \int_0^1\frac{2zdz}{[az+b(1-z)]^3}.$$

Differentiating both sides of Eq. (1) with respect to b, we get the relation (3)

$$\frac{1}{ab^2} = \int_0^1 \frac{2(1-z)dz}{[az+b(1-z)]^3}.$$

5.3 First let us consider k_σ in the numerator of the integral. Since k_σ is a odd function, the integral vanishes and it corresponds to 0 on the right hand side of the equation.

To evaluate the integral with 1 in the numerator, we need to use the contour integration. Since $d^4k = d\omega d^3k$ and $\mathbf{k}^2 = \omega^2 - \boldsymbol{k}^2$, the factor in the denominator can be written as

$$\mathbf{k}^2 + i\epsilon - L = \omega^2 + i\epsilon - (L + \boldsymbol{k}^2).$$

Then for $\epsilon \ll L + \boldsymbol{k}^2$, there are poles at $\omega = \pm[(L+\boldsymbol{k}^2)^{1/2} - i\epsilon]$, and contour integration of ω yields

$$\int_{-\infty}^{\infty} \frac{d\omega}{\omega^2 + i\epsilon - (L+\boldsymbol{k}^2)} = 2\pi i \left\{ \frac{1}{-2(L+\boldsymbol{k}^2)^{1/2}} \right\}, \tag{5.36}$$

with the contour in the upper half-plane.

Let us differentiate both sides of Eq. (5.36) with respect to L. First let us differentiate the L.H.S. of Eq. (5.36).

$$\begin{aligned} \frac{d}{dL}\left(\frac{1}{\omega^2 + i\epsilon - (L+\boldsymbol{k}^2)}\right) &= \frac{d}{dL}\left(\frac{1}{a-L}\right) = \frac{1}{(a-L)^2} \\ &\text{where } a = \omega^2 + i\epsilon - \boldsymbol{k}^2 \\ \frac{d^2}{dL^2}\left(\frac{1}{a-L}\right) &= \frac{d}{dL}\left(\frac{1}{(a-L)^2}\right) = \frac{2}{(a-L)^3} \end{aligned}$$

Now let us differentiate twice the R.H.S. of Eq. (5.36).

$$\begin{aligned} -\pi i \frac{d}{dL}(L+\boldsymbol{k}^2)^{-1/2} &= \frac{\pi i}{2}(L+\boldsymbol{k}^2)^{-3/2} \\ -\pi i \frac{d^2}{dL^2}(L+\boldsymbol{k}^2)^{-1/2} &= \frac{\pi i}{2}\frac{d}{dL}(L+\boldsymbol{k}^2)^{-3/2} = -\frac{3\pi i}{4}(L+\boldsymbol{k}^2)^{-5/2}. \end{aligned}$$

Equating the differentiated quantities on both sides, we get

$$\int \frac{d\omega}{\{\omega^2 + i\epsilon - (L+\boldsymbol{k}^2)\}^3} = \frac{3\pi}{8i}(L+\boldsymbol{k}^2)^{-5/2}. \tag{5.37}$$

Of the four dimensional integration d^4k, we have performed the integration over the energy $d\omega$ and obtained the result (5.37). Now we have to perform the integration over three-momenta d^3k.

$$\begin{aligned} \int (L+\boldsymbol{k}^2)^{-5/2} d^3k &= 4\pi \int_0^\infty (L+\boldsymbol{k}^2)^{-5/2} k^2 dk \\ &= 4\pi \frac{k^3}{3L(L+k^2)^{3/2}}\bigg|_0^\infty = \frac{4\pi}{3L}. \end{aligned}$$

Collecting all the results, we get finally

$$\begin{aligned}\int \frac{d^4k}{(2\pi)^4(\mathbf{k}^2+i\epsilon-L)^3} &= \int \frac{d\omega d^3k}{(2\pi)^4\{\omega^2+i\epsilon-(L+\boldsymbol{k}^2)\}^3} \\ &= \frac{1}{(2\pi)^4}\int \frac{3\pi}{8i}(L+\boldsymbol{k}^2)^{-5/2}d^3k \\ &= \frac{1}{(2\pi)^4}\frac{3\pi}{8i}\frac{4\pi}{3L} = \frac{1}{32\pi^2 iL} \end{aligned} \tag{5.38}$$

5.4 Let us start from the relation

$$\int_{-\infty}^{\infty} \frac{(1,k_\sigma)d^4k}{(2\pi)^4(\mathbf{k}^2+i\epsilon-L)^3} = \frac{(1,0)}{32\pi^2 iL},$$

Change the variable $\mathbf{k}\to\mathbf{k}'=\mathbf{k}-\mathbf{p}$. Then $\mathbf{k}'^2=\mathbf{k}^2+\mathbf{p}^2-2\mathbf{p}\cdot\mathbf{k}$.

$$\begin{aligned}\int \frac{d^4k'}{(2\pi)^4}\frac{1}{(\mathbf{k}'^2+i\epsilon-L)^3} &= \int \frac{d^4k'}{(2\pi)^4}\frac{1}{\{\mathbf{k}^2-2\mathbf{p}\cdot\mathbf{k}+i\epsilon-(L-\mathbf{p}^2)\}^3} \\ &= \int \frac{d^4k'}{(2\pi)^4}\frac{1}{\{\mathbf{k}^2-2\mathbf{p}\cdot\mathbf{k}+i\epsilon-\Delta\}^3} \qquad (5.39) \\ &\quad \text{where } \Delta = L-\mathbf{p}^2 \\ &= \int \frac{d^4k'}{(2\pi)^4}\frac{1}{\{\mathbf{k}'^2+i\epsilon-(\Delta+\mathbf{p}^2)\}^3} \\ &= \frac{1}{32\pi^2 i(\Delta+\mathbf{p}^2)}. \qquad (5.40)\end{aligned}$$

Now, let us consider the other integrand.

$$\int \frac{d^4k'}{(2\pi)^4}\frac{k_\sigma}{\{\mathbf{k}'^2+i\epsilon-(\Delta+\mathbf{p}^2)\}^3} = \int \frac{d^4k'}{(2\pi)^4}\frac{k'_\sigma+p_\sigma}{\{\mathbf{k}'^2+i\epsilon-(\Delta+\mathbf{p}^2)\}^3}.$$

The term k'_σ will yield zero since it is an odd function in $\mathbf{k}$. So, the term p_σ alone will contribute to the integral. Hence

$$\int \frac{d^4k'}{(2\pi)^4}\frac{k'_\sigma+p_\sigma}{\{\mathbf{k}'^2+i\epsilon-(\Delta+\mathbf{p}^2)\}^3} = \frac{p_\sigma}{32\pi^2 i(\Delta+\mathbf{p}^2)}. \tag{5.41}$$

Combining Eqs. (5.40) and (5.41) and changing the dummy variable d^4k' to d^4k, we get

$$\int \frac{d^4k}{(2\pi)^4}\frac{1;k_\sigma+p_\sigma}{\{\mathbf{k}^2+i\epsilon-(\Delta+\mathbf{p}^2)\}^3} = \frac{1;p_\sigma}{32\pi^2 i(\Delta+\mathbf{p}^2)}. \tag{5.42}$$

5.5 To evaluate the given definite integrals, we need to use the following standard integrals.

$$\int \ln x\,dx = x\ln x - x; \qquad \int x\,\ln x\,dx = \frac{x^2}{2}\ln x - \frac{x^2}{4}.$$

One can easily check the above relations by differentiating the R.H.S. and checking whether it yields the integrand on the L.H.S.

(a) Let us first expand $\ln\left(\frac{1-x}{x^2}\right)$.

$$\ln\left(\frac{1-x}{x^2}\right) = \ln(1-x) - \ln x^2 = \ln(1-x) - 2\ln x.$$

Then

$$\begin{aligned}
\int_0^1 \ln\left(\frac{1-x}{x^2}\right) dx &= \int_0^1 \ln(1-x)dx - 2\int_0^1 \ln x\, dx. \\
\int_0^1 \ln(1-x)dx &= -\int_1^0 \ln y\, dy, \\
&\quad \text{substituting } y = 1-x \text{ and } dy = -dx \\
&= -|y\ln y - y|_1^0 = -1. \\
\int_0^1 \ln x\, dx &= |1(x\ln x - x|_0^1 = 1.
\end{aligned}$$

Collecting the above results, we get

$$\int_0^1 \ln\left(\frac{1-x}{x^2}\right) dx = 1.$$

(b) A similar procedure is adopted.

$$\begin{aligned}
\int_0^1 x\ln\left(\frac{1-x}{x^2}\right) dx &= \int \{x\ln(1-x) - x\ln x^2\} dx \\
&= \int x\ln(1-x)dx - 2\int x\,\ln x\, dx \\
\int_0^1 x\ln(1-x)dx &= -\int (1-y)\ln y\, dy, \\
&\quad \text{substituting } y = 1-x \text{ and } dy = -dx \\
&= -\int_1^0 \ln y\, dy + \int_1^0 y\ln y\, dy \\
&= -|y\ln y - y|_1^0 + \left|(x^2/2)\ln x - (x^2/4)\right|_0^1 \\
&= -1 + \frac{1}{4} = -\frac{3}{4} \\
\int_0^1 x\,\ln x\, dx &= \left|\frac{x^2}{2}\ln x - \frac{x^2}{4}\right|_0^1 = -\frac{1}{4}.
\end{aligned}$$

Collecting the above results, we get

$$\int_0^1 x\ln\left(\frac{1-x}{x^2}\right) dx = -\frac{1}{4}.$$

Chapter 6

Elements of Quantum Field Theory

The formulation of Quantum Mechanics is based on the Hamiltonian theory of classical mechanics by treating the canonically conjugate dynamical variables, position and momentum as operators, obeying certain commutation relations. This is known as *the first quantization.* The alternative Lagrangian formulation of classical mechanics and the Hamiltonian action principle can be considered more fundamental than the Newtonian dynamics and they lead to the formulation of both classical and quantum field theory. The Lagrangian method can be extended to relativistic fields since the action function is relativistically invariant.

A field is specified by a number of functions, say n, of space-time.

$$\psi_\rho(x_1, x_2, x_3, t), \qquad \rho = 1, 2, 3, \ldots, n.$$

The classic example is that of electromagnetic field, which is described by the six components of the electric and magnetic field strengths at each point in space-time. A great merit of field theory is that the special theory of relativity can be incorporated in it in a simple way. Let us replace the time variable t by

$$x_0 = ct,$$

so that all the coordinates x_0, x_1, x_2, x_3 now have the dimension of length. The dynamical character of the field will be described by the field equations. They will be assumed to be partial differential equations of order, not greater than two. It will be further assumed that they can be derived

by a variational principle, in the same way as the Lagrangian equations of motion for particles are obtained from the action principle.

Let us start with a brief review of classical mechanics[1] since most of the concepts of field theory have been drawn from the Lagrangian-Hamiltonian formalism.

6.1 A Brief Review of Classical Mechanics

6.1.1 The Lagrangian-Hamiltonian formalism

Let us recapitulate the derivation of the Lagrange equations of motion from the action principle. Let $q_k(t)$, $k = 1, 2, \ldots, n$ be the generalized coordinates and $\dot{q}_k(t) = \frac{dq_k(t)}{dt}$, the generalized velocities of a mechanical system and L, its Lagrangian.

$$L = L(q_k(t), \dot{q}_k(t), t),$$

The action integral is

$$\mathscr{A} = \int_{t_1}^{t_2} L(q_k, \dot{q}_k, t)dt. \tag{6.1}$$

Now consider the variation of $q_k(t)$,

$$q_k(t) \longrightarrow q_k(t) + \delta q_k(t),$$

where $\delta q_k(t)$ is an arbitrary infinitesimal change, which vanishes at the end-points t_1 and t_2.

$$\delta q_k(t_1) = \delta q_k(t_2) = 0.$$

This variation results in the change of $\dot{q}_k(t)$ to

$$\dot{q}_k(t) + \frac{d}{dt}\delta q_k(t).$$

The variational principle[2] asserts that the action integral has a stationary value and the variation $\delta\mathscr{A}$ of $\mathscr{A}$ due to small variation in path with fixed end-points is identically zero.

$$\delta\mathscr{A} = \int_{t_1}^{t_2} \sum_k \left\{ \frac{\partial L}{\partial q_k}\delta q_k + \frac{\partial L}{\partial \dot{q}_k}\frac{d}{dt}\delta q_k \right\} dt = 0. \tag{6.2}$$

[1]For greater details, the reader is referred to: H. Goldstein, *Classical Mechanics*, Narosa Publishing House, New Delhi (Indian Edition) (2001).

[2]Of the various paths available between the two given end points, the system chooses the path for which the action integral $\mathscr{A}$ is a minimum. This is found by the method of *calculus of variations*.

The second term in the integral is evaluated using the principle of integration by parts.

$$\begin{aligned}\int \frac{\partial L}{\partial \dot{q}_k}\frac{d}{dt}\delta q_k dt &= \frac{\partial L}{\partial \dot{q}_k}\delta q_k\Big|_{t_1}^{t_2} - \int \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}\delta q_k dt \\ &= -\int \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}\delta q_k dt, \end{aligned} \tag{6.3}$$

since $\frac{\partial L}{\partial \dot{q}_k}\delta q_k\Big|_{t_1}^{t_2}$ vanishes because $\delta q_k(t_1) = \delta q_k(t_2) = 0$, due to the fixed end-points.

Substituting (6.3) into Eq. (6.2), we get

$$\delta\mathscr{A} = \int_{t_1}^{t_2}\sum_k \left\{\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}\right\}\delta q_k dt = 0. \tag{6.4}$$

Since the generalized coordinates q_k are independent and the variations δq_k are arbitrary, the condition that $\delta\mathscr{A} = 0$ requires that the coefficients of δq_k in the integrand of (6.4) should separately vanish.

$$\frac{\partial L}{\partial q_k} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k} = 0, \qquad k = 1, 2, \ldots, n. \tag{6.5}$$

Equation (6.5) is known as *the Euler-Lagrange equation of motion.*

The momentum p_k conjugate to the coordinate q_k is given by

$$p_k = \frac{\partial L}{\partial \dot{q}_k}. \tag{6.6}$$

Differentiating p_k with respect to t and using the relation (6.5), we obtain

$$\dot{p}_k = \frac{\partial L}{\partial q_k}. \tag{6.7}$$

If T is the kinetic energy and V is the potential energy, then the Lagrangian of the system is given by

$$L = T - V, \tag{6.8}$$

and the Hamiltonian function is given by

$$H(q_k, p_k, t) = \sum_k p_k\dot{q}_k - L. \tag{6.9}$$

It follows that

$$\frac{\partial H}{\partial p_k} = \dot{q}_k \qquad \text{and} \qquad \frac{\partial H}{\partial q_k} = -\frac{\partial L}{\partial q_k}.$$

Using the Euler-Lagrange Eq. (6.5) and the relation (6.6), we deduce *Hamilton's canonical equations of motion.*

$$\dot{q}_k = \frac{\partial H}{\partial p_k}; \qquad \dot{p}_k = -\frac{\partial H}{\partial q_k}. \tag{6.10}$$

From (6.10), one can deduce the equations of motion for a general dynamical variable $F \equiv F(q_k, p_k, t)$.

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial t} + \sum_k \left(\frac{\partial F}{\partial q_k}\frac{\partial q_k}{\partial t} + \frac{\partial F}{\partial p_k}\frac{\partial p_k}{\partial t} \right) \\ &= \frac{\partial F}{\partial t} + [F, H]_{\mathrm{PB}}, \end{aligned} \tag{6.11}$$

where

$$[F, H]_{\mathrm{PB}} = \sum_k \left(\frac{\partial F}{\partial q_k}\frac{\partial H}{\partial p_k} - \frac{\partial H}{\partial q_k}\frac{\partial F}{\partial p_k} \right) \tag{6.12}$$

is *the Poisson Bracket.*

6.1.2 The classical fields

From a discrete mechanical system, it is possible to go to a continuous system when the number of degrees of freedom becomes infinite. Consider a classical field $\psi(\boldsymbol{x}, t)$ which is continuous and described by an infinite set of parameters $\boldsymbol{x}$ and t. The Lagrangian formalism can be extended to the fields and the variational principle can be incorporated in an analogous manner by introducing certain modifications.

1. Since the field $\psi(\boldsymbol{x}, t)$ depends on the continuous variable $\boldsymbol{x}$, we have to invoke the concept of *Lagrangian density.*

2. The theory can be made relativistically invariant by treating both $\boldsymbol{x}$ and t coordinates on equal footing. This can be done by using various metrics but we find it more convenient to use the metric

$1, -1, -1, -1$ instead of the Minkowski space[3], with the time coordinate $x_0 = ct$. By this procedure, we make all the four coordinates of the same dimension of length.

Let us start with the Lagrangian density $\mathscr{L}$.

$$\mathscr{L}\left(\psi_\rho, \frac{\partial \psi_\rho}{\partial x_\mu}, \mathbf{x}\right) \equiv \mathscr{L}\left(\psi_\rho, \psi_{\rho,\mu}, \mathbf{x}\right),$$

where, for brevity, a notation $\psi_{\rho,\mu}$ is used to denote $\frac{\partial \psi_\rho}{\partial x_\mu}$. The Lagrangian density is a function of field functions ψ_ρ, its first order derivatives and the space-time coordinates $(\mathbf{x} : x_0 = ct, x_1, x_2, x_3)$. It is important that $\mathscr{L}$ should not depend on second and higher order derivatives of ψ_ρ.

The Lagrangian and the action integral are given by

$$L = \int \mathscr{L}\left(\psi_\rho, \psi_{\rho,\mu}, \mathbf{x}\right)\, d^3x; \tag{6.13}$$

$$\mathscr{A} = \int_V \mathscr{L}\left(\psi_\rho, \psi_{\rho,\mu}, \mathbf{x}\right)\, d^4x; \tag{6.14}$$

where

$$d^4x = dx_0 dx_1 dx_2 dx_3 = cdtdxdydz,$$

and V is a certain domain in space-time, bounded by a hypersurface S. According to the principle of least action, the variation $\delta\mathscr{A}$ in $\mathscr{A}$ for arbitrary domains V is zero for the variations of $\psi_\rho(\mathbf{x})$.

$$\psi_\rho(\mathbf{x}) \longrightarrow \psi_\rho(\mathbf{x}) + \delta\psi_\rho(\mathbf{x}),$$

where $\delta\psi_\rho(\mathbf{x})$ is an arbitrary infinitesimal variation of the first order, which is continuous and differentiable and vanishes on the boundary S of V, *i.e.*.

$$\delta\psi_\rho(\mathbf{x}) = 0, \qquad \text{on the surface } S \text{ of } V. \tag{6.15}$$

[3]The Minkowski space is a natural extension of the ordinary three dimensional Cartesian space with which we are familiar, to the four dimensional space-time which is required for any theory that includes the special theory of relativity. This avoids the use of metric tensor and the artificial distinction between the covariant and contravariant quantities. However this is not possible for the treatment of the general theory of relativity with the concept of curved space. So, many authors prefer to use the metric $1, -1, -1, -1$ with the accompanying distinction between covariant and contravariant quantities.

This causes a change in the derivative.

$$\frac{\partial\psi_\rho}{\partial x_\mu} \longrightarrow \frac{\partial\psi_\rho}{\partial x_\mu} + \frac{\partial}{\partial x_\mu}\delta\psi_\rho.$$

Thus

$$\delta\mathscr{A} = \int_V \sum_\rho \left\{ \frac{\partial\mathscr{L}}{\partial\psi_\rho}\delta\psi_\rho + \sum_\mu \frac{\partial\mathscr{L}}{\partial\frac{\partial\psi_\rho}{\partial x_\mu}} \frac{\partial}{\partial x_\mu}\delta\psi_\rho \right\} d^4x = 0. \tag{6.16}$$

Integrating the second term in the curly brackets by parts, we get

$$\int_V \frac{\partial\mathscr{L}}{\partial\frac{\partial\psi_\rho}{\partial x_\mu}} \frac{\partial}{\partial x_\mu}\delta\psi_\rho d^4x = \left.\frac{\partial\mathscr{L}}{\partial\frac{\partial\psi_\rho}{\partial x_\mu}}\delta\psi_\rho\right|_S - \int_V \frac{\partial}{\partial x_\mu}\frac{\partial\mathscr{L}}{\partial\frac{\partial\psi_\rho}{\partial x_\mu}}\delta\psi_\rho d^4x. \tag{6.17}$$

The first term on the right hand side of Eq. (6.17) vanishes because of the boundary condition (6.15) and hence the variational principle leads to the result

$$\delta\mathscr{A} = \int_V \sum_\rho \left\{ \frac{\partial\mathscr{L}}{\partial\psi_\rho} - \sum_\mu \frac{\partial}{\partial x_\mu}\frac{\partial\mathscr{L}}{\partial\frac{\partial\psi_\rho}{\partial x_\mu}} \right\} \delta\psi_\rho d^4x = 0. \tag{6.18}$$

Since this equation holds for arbitrary $\delta\psi_\rho$ and also valid for arbitrary domain V, the resulting relations

$$\frac{\partial\mathscr{L}}{\partial\psi_\rho} - \sum_\mu \frac{\partial}{\partial x_\mu}\frac{\partial\mathscr{L}}{\partial\frac{\partial\psi_\rho}{\partial x_\mu}} = 0, \qquad \rho = 1, 2, \ldots, n, \tag{6.19}$$

which are analogous to the Euler-Lagrange equations (6.5) are obtained. These equations, in turn, lead to the field equations. Since $\mathscr{L}$ does not involve derivatives of ψ_ρ of order higher than the first, the resulting field equations will be utmost of second order.

One could, in principle, consider field equations that are not derivable from a variational principle, just as in mechanics we can have systems with frictional and dissipative forces for which equations of motion cannot be so obtained. Non-conservative forces, like friction, appear at a macroscopic level since we wish to neglect microscopic complications. So, at the microscopic level, the fundamental laws can be put in the form of the principle of least action.

Using the short-hand notation, the Euler-Lagrange Eq. (6.19) can be rewritten as

$$\frac{\partial \mathscr{L}}{\partial \psi_\rho} - \partial_\mu \frac{\partial \mathscr{L}}{\partial(\partial_\mu \psi_\rho)} = 0, \qquad \rho = 1, 2, \ldots, n, \tag{6.20}$$

with the convention that a summation is to be made on the repeated indices.

The conjugate momentum density $\Pi(\mathbf{x})$ and the Hamiltonian density $\mathscr{H}(\mathbf{x})$ are obtained in a way, analogous to Eqs. (6.6) and (6.9).

$$\Pi(\mathbf{x}) = \frac{\partial \mathscr{L}}{\partial \dot{\psi}}. \tag{6.21}$$

$$\mathscr{H}(\mathbf{x}) = \Pi(\mathbf{x})\dot{\psi}(\mathbf{x}) - \mathscr{L}. \tag{6.22}$$

6.2 Quantization of the Field

The transition from the classical field theory to the quantum field theory is made by postulating the field variables ψ and Π as field operators and prescribing certain algebraic relations between them. The prescription of these algebraic relations (commutators or anticommutators) are known as *the second quantization.*

It may be recalled that some of the field equations, for example, the Schrödinger equation, the Klein-Gordon equation and the Dirac equation were obtained earlier by treating the dynamical variables $\boldsymbol{x}$ and $\boldsymbol{p}$ as operators obeying certain commutation relations. That is known as *the first quantization.* Since in quantum field theory, the field functions are treated as operators and certain algebraic relations are postulated between them for the second time, this is known as the second quantization.

The quantization of the field results in the description of the field in terms of "particles" or more precisely, "the field quanta" which are the carriers of energy, momentum and charge of the field. Let us illustrate the method of quantization by considering the Schrödinger field. It is a non-relativistic field and it allows the quantization procedure by choosing either the commutation relations or anticommutation relations between the field operators. Later, we shall see that in the case of relativistic fields such as the Klein-Gordon field or the Dirac field, only one of them is admissible.

6.2.1 The Schrödinger field

Let us consider the Schrödinger field which satisfies the field equation

$$i\hbar\frac{d\psi}{dt} + \frac{\hbar^2}{2m}\nabla^2\psi - V\psi = 0. \tag{6.23}$$

Quantum mechanics treats Eq. (6.23) as the equation of motion of a particle of mass m moving in an external potential V. Here, we treat it as a classical field equation which, when quantized, will yield an assembly of a large number indistinguishable particles. Since this is the second time that quantization is done, this is known as the second quantization. It may be recalled that Eq. (6.23) has been obtained earlier by treating the dynamical variables $\boldsymbol{x}$ and $\boldsymbol{p}$ as operators obeying the commutation relation $[x, p_x]_- = i\hbar$; which is known as the first quantization.

The Lagrangian density for the Schrödinger field is given by

$$\mathscr{L}_S = i\hbar\psi^*\dot{\psi} - \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^* \cdot \boldsymbol{\nabla}\psi - V\psi^*\psi, \tag{6.24}$$

which, when substituted in the Euler-Lagrange Equation

$$\frac{\partial\mathscr{L}}{\partial\psi} - \boldsymbol{\nabla}\cdot\frac{\partial\mathscr{L}}{\partial(\boldsymbol{\nabla}\psi)} - \frac{\partial}{\partial t}\left(\frac{\partial\mathscr{L}}{\partial\dot{\psi}}\right) = 0 \tag{6.25}$$

yields Eq. (6.23). Equation (6.25) is the same as Eq. (6.20) but given in the expanded form, separating the space and time components.

The conjugate field $\Pi(\boldsymbol{x}, t)$ is given by

$$\Pi(\boldsymbol{x}, t) = \frac{\partial\mathscr{L}_S}{\partial\dot{\psi}} = i\hbar\psi^*(\boldsymbol{x}, t). \tag{6.26}$$

The Hamiltonian density $\mathscr{H}$ and the Hamiltonian H are given by

$$\mathscr{H} = \Pi\dot{\psi} - \mathscr{L}_S = \frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^* \cdot \boldsymbol{\nabla}\psi - V\psi^*\psi : \tag{6.27}$$

$$H = \int_V \mathscr{H} d^3x = \int_V \left(\frac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^* \cdot \boldsymbol{\nabla}\psi - V\psi^*\psi\right) d^3x, \tag{6.28}$$

where the suffix V on the integral denotes the volume of integration, over which the field extends.

Quantization

The quantization of the Schrödinger field can be done by expanding the field operator ψ in terms of the complete set of eigenfuntions $u_k(\boldsymbol{x})$ of the field equation (6.23). This is known as the Fourier decomposition of the field into its normal modes.

$$\psi(\boldsymbol{x},t) = \sum_k a_k(t) u_k(\boldsymbol{x}), \tag{6.29}$$

where $a_k(t)$ is the Fourier coefficient which is interpreted as the annihilation operator of the field quantum with momentum $\boldsymbol{k}$. The functions $u_k(\boldsymbol{x})$ satisfy the following relations:

$$\int u_k^*(\boldsymbol{x}) u_l(\boldsymbol{x}) d^3x = \delta_{kl}; \tag{6.30}$$

$$\sum_k u_k(\boldsymbol{x}) u_k^*(\boldsymbol{x}') = \delta(\boldsymbol{x}-\boldsymbol{x}'); \tag{6.31}$$

and they are the single particle eigenfunctions of the Hamiltonian

$$H_P = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2 + V,$$

with eigenvalue ϵ_k.

$$H_P u_k(\boldsymbol{x}) = \epsilon_k u_k(\boldsymbol{x}).$$

The Hermitian conjugate of ψ is

$$\psi^\dagger(\boldsymbol{x},t) = \sum_k a_k^\dagger(t) u_k^*(\boldsymbol{x}).$$

The quantization is done by postulating certain algebraic relations between the field operators ψ and $\psi^\dagger$ or alternatively between the operators $a_k(t)$ and $a_l^\dagger(t)$.

In case of the Schrödinger field, which is a non-relativistic field, it is possible to choose commutation relations or anticommutation relations[4] between a_k and $a_k^\dagger$. The first one corresponds to a system of Bosons and the second one corresponds to a system of Fermions.

[4]The square brackets are used to denote the commutation relation and the curly brackets denote the anticommutation relation.

$$[a_k, a_l^\dagger]_- = a_k a_l^\dagger - a_l^\dagger a_k = \delta_{kl}; \qquad \{a_k, a_l^\dagger\}_+ = a_k a_l^\dagger + a_l^\dagger a_k = \delta_{kl}.$$

6.2.2 Quantization into Bosons

Let us postulate the following commutation rules for the field operator $\psi(\boldsymbol{x}, t)$ and its conjugate momentum $\Pi(\boldsymbol{x}, t)$ defined by Eq. (6.26):

$$[\psi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)]_- = i\hbar\delta(\boldsymbol{x} - \boldsymbol{x}') \tag{6.32}$$

$$[\psi(\boldsymbol{x}, t), \psi(\boldsymbol{x}', t)]_- = 0 = [\Pi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)]_- \tag{6.33}$$

This is analogous to the Heisenberg commutation relations used in Quantum Mechanics for $\boldsymbol{x}$ and $\boldsymbol{p}$. This, in turn, leads to the commutation relations between the Fourier coefficients a_k, $a_k^\dagger$ on using the expansion (6.29).

$$[a_k, a_l^\dagger]_- = \delta_{kl}; \tag{6.34}$$

$$[a_k, a_l]_- = [a_k^\dagger, a_l^\dagger]_- = 0; \tag{6.35}$$

where all the operators refer to the same time. The operators a_k and $a_k^\dagger$ are called the annihilation and creation operators of the field quanta. The entire description of the system – its state vector, the number and energy of the system and the effects of the operators on the system – can hereafter be studied in terms of these operators.

1. The number operator N_k is an Hermitian operator

$$N_k = a_k^\dagger a_k, \tag{6.36}$$

with eigenvalues $n_k = 0, 1, \ldots, \infty$.

2. The state vector is better described in the occupation number representation, better known as the Fock space, giving the number of quanta in each state.

$$|n_1, n_2, \ldots, n_k, \ldots\rangle = c\,(a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2}\ldots(a_k^\dagger)^{n_k}\ldots|0\rangle,$$

with

$$c = \frac{1}{(n_1!n_2!\ldots n_k!\ldots)^{1/2}}.$$

The vacuum state $|0\rangle$ is defined by

$$N_k|0\rangle = 0, \qquad \text{for all } k.$$

3. The effect of these operators on the state vector is given below:

$$a_k \left|n_1, n_2, \ldots, n_k, \ldots\right\rangle = \sqrt{n_k}\, \left|n_1, n_2, \ldots, n_k - 1, \ldots\right\rangle; \quad (6.37)$$

$$a_k^\dagger \left|n_1, n_2, \ldots, n_k, \ldots\right\rangle = \sqrt{n_k + 1}\, \left|n_1, n_2, \ldots, n_k + 1, \ldots\right\rangle; \quad (6.38)$$

$$N_k \left|n_1, n_2, \ldots, n_k, \ldots\right\rangle = n_k \left|n_1, n_2, \ldots, n_k, \ldots\right\rangle. \quad (6.39)$$

4. The Hamiltonian of the system and its energy are given by

$$H = \sum_{kl} a_k^\dagger a_l \int_V \left(\frac{\hbar^2}{2m} \boldsymbol{\nabla} u_k^* \cdot \boldsymbol{\nabla} u_l + V u_k^* u_l\right) d^3x \quad (6.40)$$

$$= \sum_k N_k \epsilon_k. \quad (6.41)$$

The total energy of the field in the state $|n_1, n_2, \ldots, n_k, \ldots\rangle$ is given by

$$E = \sum_k n_k \epsilon_k.$$

Since the given state u_k can be occupied by any number of particles, the field represents an assembly of Bosons.

6.2.3 Quantization into Fermions

For a system of Fermions, the exclusion principle restricts the occupation number n_k of a particular state to 0 or 1. It has been shown by Jordan and Wigner[5] that this condition would be satisfied if the commutation relations (6.34) and (6.35) are replaced by the anticommutation relations given below.

$$\{a_k, a_l^\dagger\}_+ = \delta_{kl}; \quad (6.42)$$

$$\{a_k, a_l\}_+ = 0 = \{a_k^\dagger, a_l^\dagger\}_+. \quad (6.43)$$

All the operators refer to the same time. Besides, from Eq. (6.43), we obtain

$$a_k a_k = 0 = a_k^\dagger a_k^\dagger, \quad (6.44)$$

such that

$$N_k^2 = a_k^\dagger a_k a_k^\dagger a_k = a_k^\dagger (1 - a_k^\dagger a_k) a_k = N_k, \quad (6.45)$$

[5]P. Jordan and E. P. Wigner, *Z. Physik*, **47**, 631 (1928).

which, in turn, yields the relation

$$N_k(N_k - 1) = 0. \tag{6.46}$$

This condition restricts the eigenvalue n_k of N_k to

$$n_k = 0, 1. \tag{6.47}$$

The following results follow from the above relations.

$$\begin{aligned}
|n_1, n_2, \ldots, n_k, \ldots\rangle &= (a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2} \ldots (a_k^\dagger)^{n_k} \ldots |0\rangle; && (6.48)\\
a_k |n_1, n_2, \ldots, n_k, \ldots\rangle &= (-1)^{S_k} n_k |n_1, n_2, \ldots, n_k - 1, \ldots\rangle ; && (6.49)\\
a_k^\dagger |n_1, n_2, \ldots, n_k, \ldots\rangle &= (-1)^{S_k} (1 - n_k) |n_1, n_2, \ldots, n_k + 1, \ldots\rangle ; && (6.50)\\
N_k |n_1, n_2, \ldots, n_k, \ldots\rangle &= n_k |n_1, n_2, \ldots, n_k, \ldots\rangle ; && (6.51)
\end{aligned}$$

with

$$S_k = \sum_{r=1}^{k-1} n_r.$$

Eqs. (6.49) and (6.50) state that the annihilation operator acting on an empty state or the creation operator acting on a filled state yields zero.

Since, for the Fermions, n_k can assume only two values 0 and 1, a 2×2 matrix representation can be given to the occupation number operator N_k.

$$N_k = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Consequently, the matrix representation for the annihilation and creation operators become

$$a_k = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \qquad a_k^\dagger = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

The Hamiltonian, the total energy and the total number of particles of a system of Fermions that the field represents are given by

$$H = \sum_k N_k \epsilon_k; \qquad E = \sum_k n_k \epsilon_k; \qquad N = \sum_k n_k.$$

Since the total number operator $N = \sum_k N_k$ commutes with the Hamiltonian,

$$[N, H] = \sum_{kl} [N_k, N_l] \epsilon_l = 0,$$

it follows that the total number of Fermions in the field is conserved.

It can be shown that the anticommutation relations between the annihilation and creation operators for the Fermion field arise from the anticommutation relations between the field operator $\psi(\boldsymbol{x}, t)$ and its conjugate momentum $\Pi(\boldsymbol{x}, t)$.

$$\{\psi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)\}_+ = i\hbar\delta(\boldsymbol{x} - \boldsymbol{x}') \tag{6.52}$$

$$\{\psi(\boldsymbol{x}, t), \psi(\boldsymbol{x}', t)\}_+ = 0 = \{\Pi(\boldsymbol{x}, t), \Pi(\boldsymbol{x}', t)\}_+ \tag{6.53}$$

This anticommutation relations have no classical analogue.

We can also represent the particle number operator N in terms of the field variable $\psi(\boldsymbol{x}, t)$, using the relation (6.29).

$$\begin{aligned} N &= \sum_k a_k^\dagger a_k = \sum_k \int_V u_k(\boldsymbol{x})\psi^\dagger(\boldsymbol{x}, t)u_k^*(\boldsymbol{x}')\psi(\boldsymbol{x}', t)d^3x d^3x' \\ &= \int_V \delta(\boldsymbol{x} - \boldsymbol{x}')\psi^\dagger(\boldsymbol{x}, t)\psi(\boldsymbol{x}', t)d^3x d^3x', \quad \text{summing over } \boldsymbol{k} \\ &= \int_V \psi^\dagger(\boldsymbol{x}, t)\psi(\boldsymbol{x}, t)d^3x. \end{aligned} \tag{6.54}$$

6.3 Relativistic Fields

In particle mechanics, the space and time coordinates play different roles. The space coordinates are the mechanical variables whereas the time is a parameter. But in the case of classical fields, the space and time coordinates are treated similarly and they are the parameters specifying the space-time continuum, with which the field variables are defined. This facilitates the relativistically covariant formulation of the classical fields which are commonly known as relativistic fields that are manifestly Lorentz covariant.

With the introduction of the fourth coordinate $x_0 = ct$ in the place of time coordinate, the action integral $\mathscr{A}$ defined by Eq. (6.14) changes only by a multiplicative factor but does not affect the formulation of the Hamilton's action principle which yields the same Lagrange equations of motion (6.19). Please note that the only term

$$\frac{\partial}{\partial x_\mu} \frac{\partial \mathscr{L}}{\partial \left(\frac{\partial \mathscr{L}}{\partial x_\mu}\right)},$$

that depends on x_μ in Eq. (6.19) is not affected by the change of scale of any x_μ.

In relativistic field theory, we should expect that all the quantities that define the field and the associated equations should be covariant under Lorentz transformation. This means that the quantities should be either world scalars or world vectors or world tensors. The action integral $\mathscr{A}$ is a world scalar. Similarly, the Lagrangian density $\mathscr{L}$ and the Hamiltonian density $\mathscr{H}$ are world scalars. The four-dimensional volume element d^4x is invariant under Lorentz transformation. Let us show that the second rank tensor constructed from the Lagrangian density using the Lagrange equations of motion is a world tensor of second rank, the components of which yields the momentum and energy density of the field.

Given the Lagrangian density

$$\mathscr{L}(\psi, \frac{\partial\psi}{\partial x_\mu}, x_\mu),$$

the total derivative of the Lagrangian density is given by

$$\frac{d\mathscr{L}}{dx_\mu} = \frac{\partial\mathscr{L}}{\partial\psi}\frac{\partial\psi}{\partial x_\mu} + \frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_\nu}\right)}\frac{\partial}{\partial x_\mu}\left(\frac{\partial\psi}{\partial x_\nu}\right) + \frac{\partial\mathscr{L}}{\partial x_\mu}. \tag{6.55}$$

From Euler-Lagrange equation (6.19), we obtain

$$\frac{\partial\mathscr{L}}{\partial\psi} = \frac{\partial}{\partial x_\nu}\frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_\nu}\right)}$$

. Substituting this into Eq. (6.55), we get

$$\begin{aligned}\frac{d\mathscr{L}}{dx_\mu} &= \frac{\partial}{\partial x_\nu}\frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_\nu}\right)}\frac{\partial\psi}{\partial x_\mu} + \frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_\nu}\right)}\frac{\partial}{\partial x_\mu}\left(\frac{\partial\psi}{\partial x_\nu}\right) + \frac{\partial\mathscr{L}}{\partial x_\mu} \\ &= \frac{d}{dx_\nu}\left\{\frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_\nu}\right)}\frac{\partial\psi}{\partial x_\mu}\right\} + \frac{\partial\mathscr{L}}{\partial x_\mu}. \end{aligned} \tag{6.56}$$

On rearrangement, we get

$$\frac{d}{dx_\nu}\left\{\frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_\nu}\right)}\frac{\partial\psi}{\partial x_\mu} - \mathscr{L}\delta_{\mu\nu}\right\} = -\frac{\partial\mathscr{L}}{\partial x_\mu}. \tag{6.57}$$

The quantity within the curly bracket is a second rank tensor, denoted by the symbol $T_{\mu\nu}$.

$$T_{\mu\nu} = \frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_\nu}\right)}\frac{\partial\psi}{\partial x_\mu} - \mathscr{L}\delta_{\mu\nu}. \tag{6.58}$$

If the Lagrangian density $\mathscr{L}$ does not depend explicitly on x_μ, *i.e.* if

$$\frac{\partial \mathscr{L}}{\partial x_\mu} = 0,$$

then $\mathscr{L}$ represents a free field. There are no sources or sinks or any interaction with other fields or particles. It follows from Eq. (6.57) that for a free field

$$\sum_\nu \frac{dT_{\mu\nu}}{dx_\nu} = 0. \tag{6.59}$$

It can easily be verified that

$$\begin{aligned} T_{44} &= \frac{\partial \mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_4}\right)}\frac{\partial\psi}{\partial x_4} - \mathscr{L} \\ &= \frac{\partial \mathscr{L}}{\partial \dot{\psi}}\dot{\psi} - \mathscr{L}. \end{aligned} \tag{6.60}$$

This is the energy density $\mathscr{H}(\mathbf{x})$ of the free field as can be seen from Eqs. (6.21) and (6.22).

The other components of the tensor are

$$\begin{aligned} T_{k4} &= \frac{\partial \mathscr{L}}{\partial\left(\frac{\partial\psi}{\partial x_4}\right)}\frac{\partial\psi}{\partial x_k}, \quad (k = 1, 2, 3) \\ &= ic\frac{\partial \mathscr{L}}{\partial \dot{\psi}}\frac{\partial\psi}{\partial x_k} \\ &= ic\Pi\frac{\partial\psi}{\partial x_k}, \quad \text{using Eq. (6.21).} \end{aligned} \tag{6.61}$$

Defining the momentum density $\mathscr{P}$ by

$$\mathscr{P}_k = \Pi\frac{\partial\psi}{\partial x_k}, \tag{6.62}$$

we find that

$$T_{k4} = ic\mathscr{P}_k. \tag{6.63}$$

Review Questions

6.1 Discuss briefly the Lagrangian-Hamiltonian formalism in classical mechanics and deduce the Euler-Lagrange equations of motion and also Hamilton's canonical equations of motion. Explain how these concepts can be applied to the classical fields.

6.2 What is meant by the second quantization? Discuss how the Schrödinger field can be quantized to yield a system of Bosons. Define the number operator and express the Hamiltonian in terms of the number operator.

6.3 Discuss how the Schrödinger equation can be quantized to yield a system of Fermions. Show that the number operator can assume only two values, 0 or 1. Obtain an expression for the Hamiltonian in terms of the number operator.

6.4 What are the features that a relativistic field should exhibit? Given the Lagrangian density, construct a second rank tensor using Euler-Lagrange equations of motion and show that the components of the second rank tensor yield the momentum and energy density of the field.

Problems

6.1 Show that in the case of motion of a single particle, the principle of least action yields the second law of Newton.

6.2 Show that the Lagrangian density for the Schrödinger field which obeys Schrödinger equation is

$$\mathscr{L}_S = i\hbar\phi^*\dot{\phi} - \frac{\hbar^2}{2m}\boldsymbol{\nabla}\phi^* \cdot \boldsymbol{\nabla}\phi - V\phi^*\phi.$$

6.3 Given the Lagrangian density

$$\mathscr{L} = \frac{1}{2}\left\{\dot{\phi}^2 - (\boldsymbol{\nabla}\phi)^2 - m^2\phi^2\right\},$$

deduce the equation for the field ϕ that the Lagrangian density represents.

Solutions to Problems

6.1 Newton's second law states that a point particle of mass m moving in a potential field $V(x)$ experiences a force F resulting in an acceleration $\frac{d^2x}{dt^2}$ of the particle.

$$m\frac{d^2x}{dt^2} = F = -\frac{dV(x)}{dt}.$$

The Lagrangian L of the particle is given by

$$L = T - V = \frac{1}{2}m\left(\frac{dx}{dt}\right)^2 - V(x),$$

where T is the kinetic energy. The action $\mathscr{A}$ is defined by

$$\mathscr{A} = \int_{t_1}^{t_2} L(x, \dot{x})dt,$$

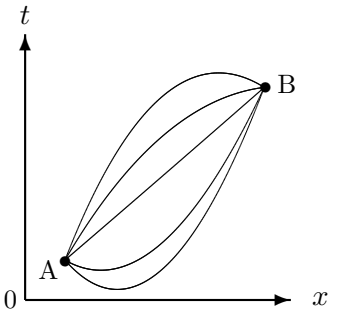


Figure 6.1: Multiple paths available for the particle between two space-time points A and B, of which the particle chooses one for which the action $\mathscr{A}$ is minimal.

where the integral is taken over the entire path of the particle from time t_1 to t_2 as shown in Fig. There is an infinite number of possible paths for 2the particle but what the actual path that the particle will take is given by the principle of least action. In other words, the particle chooses a path for which the action $\mathscr{A}$ is minimum.

Consider a variation in the path

$$x(t) \rightarrow x'(t) = x(t) + a(t), \qquad a << x.$$

Since the end points of the path are fixed,

$$a(t_1) = a(t_2) = 0.$$

On the substitution $x \rightarrow x'$, the action $\mathscr{A}$ becomes

$$\begin{aligned}
\mathscr{A} \rightarrow \mathscr{A}' &= \int_{t_1}^{t_2} \left[\tfrac{1}{2}m(\dot{x}+\dot{a})^2 - V(x+a)\right] dt \\
&= \int_{t_1}^{t_2} \left[\tfrac{1}{2}m\dot{x}^2 + m\dot{x}\dot{a} - \{V(x) - aV'(x)\}\right] dt + O(a^2) \\
&= \mathscr{A} + \int_{t_1}^{t_2} [m\dot{x}\dot{a} - aV'(x)]dt \\
&= \mathscr{A} + \delta\mathscr{A},
\end{aligned}$$

where

$$\delta\mathscr{A} = \int_{t_1}^{t_2} [m\dot{x}\dot{a} - aV'(x)]dt, \qquad V'(x) = \frac{dV(x)}{dx}.$$

If $\mathscr{A}$ is a minimum under the variation in x, then $\delta\mathscr{A} = 0$. Integrating the first term in $\delta\mathscr{A}$ by parts, we get

$$\int_{t_1}^{t_2} \dot{x}\dot{a}dt = \dot{x}a|_{t_1}^{t_2} - \int_{t_1}^{t_2} a\ddot{x}dt = -\int_{t_1}^{t_2} a\ddot{x}dt,$$

since $a(t_1) = a(t_2) = 0$ because of the fixed end points. Thus we get

$$\delta\mathscr{A} = -\int_{t_1}^{t_2} [ma\ddot{x} + aV'(x)]dt = 0,$$

which is satisfied if

$$m\ddot{x} = -V'(x).$$

Thus we have deduced Newton's second law of motion from the principle of least action.

6.2 Given the Lagrangian density $\mathscr{L}_S$ of the field, find the field equation which satisfies the Euler-Lagrange equation (6.19).

Let us write explicitly the Euler-Lagrange equation.

$$\begin{aligned} \frac{\partial\mathscr{L}}{\partial\phi} - \sum_{\mu=0}^{3} \frac{\partial}{\partial x^\mu} \frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\phi}{\partial x^\mu}\right)} &= 0 \\ \text{or} \quad \frac{\partial\mathscr{L}}{\partial\phi} - \frac{\partial}{\partial t}\frac{\partial\mathscr{L}}{\frac{\partial\phi}{\partial t}} - \sum_{k=1}^{3} \frac{\partial}{\partial x_k} \frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\phi}{\partial x^k}\right)} &= 0. \end{aligned} \tag{6.64}$$

Given the Lagrangian density

$$\mathscr{L}_s = i\hbar\phi^*\dot{\phi} - \frac{\hbar^2}{2m}\boldsymbol{\nabla}\phi^* \cdot \boldsymbol{\nabla}\phi - V\phi^8\phi,$$

we find

$$\begin{aligned} &\tfrac{\partial\mathscr{L}_S}{\partial\phi} = -V\phi^* \\ &\tfrac{\partial\mathscr{L}_S}{\partial\boldsymbol{\nabla}\phi} = -\tfrac{\hbar^2}{2m}\boldsymbol{\nabla}\phi^*, \qquad \tfrac{\partial}{\partial x^k}\left(\tfrac{\partial\mathscr{L}_S}{\partial\boldsymbol{\nabla}\phi}\right) = -\tfrac{\hbar^2}{2m}\nabla^2\phi^*. \\ &\tfrac{\partial\mathscr{L}_S}{\partial\dot{\phi}} = i\hbar\phi^*, \qquad \tfrac{\partial\mathscr{L}_S}{\partial t}\left(\tfrac{\partial\mathscr{L}_S}{\partial\dot{\phi}}\right) = i\hbar\tfrac{\partial\phi^*}{\partial t} \end{aligned} \tag{6.65}$$

Substituting (6.65) into (6.64), we get

$$-\frac{\hbar^2}{2m}\nabla^2\phi^* + V\phi^* + i\hbar\frac{\partial\phi^*}{\partial t} = 0,$$

the complex conjugate of which is the Schrödinger equation.

$$-\frac{\hbar^2}{2m}\nabla^2\phi + V\phi - i\hbar\frac{\partial\phi}{\partial t} = 0,$$

6.3 Given the Lagrangian density

$$\mathscr{L} = \frac{1}{2}\left\{\dot{\phi}^2 - (\boldsymbol{\nabla}\phi)^2 - m^2\phi^2\right\},$$

it can be written in the four-component space-time notation.

$$\mathscr{L} = \frac{1}{2}\sum_{\mu}\left(\frac{\partial\phi}{\partial x^{\mu}}\right)^{2} - \frac{1}{2}m^{2}\phi^{2}.$$

Since the Lagrangian density is given, we need to find the field equation that satisfies the Euler-Lagrangian equation.

$$\frac{\partial\mathscr{L}}{\partial\phi} - \sum_{\mu}\frac{\partial}{\partial x^{\mu}}\frac{\partial\mathscr{L}}{\partial\frac{\partial\phi}{\partial x^{\mu}}} = 0.$$

Since

$$\begin{aligned}
\frac{\partial\mathscr{L}}{\partial\phi} &= -m^{2}\phi \\
\frac{\partial\mathscr{L}}{\partial\frac{\partial\phi}{\partial x^{\mu}}} &= \frac{\partial\phi}{\partial x^{\mu}} \\
\frac{\partial}{\partial x^{\mu}}\frac{\partial\mathscr{L}}{\partial\frac{\partial\phi}{\partial x^{\mu}}} &= \frac{\partial^{2}\phi}{\partial x^{\mu 2}}
\end{aligned}$$

Substituting these values in Euler-lagrange equation, we get the field equation which can be identified as the Klein-Gordon equation

$$(\Box + m^{2})\phi(\mathbf{x}) = 0,$$

where

$$\Box = \frac{\partial}{\partial x^{\mu}}\frac{\partial}{\partial x_{\mu}} = \frac{\partial^{2}}{\partial t^{2}} - \boldsymbol{\nabla}^{2},$$

in natural units.

Chapter 7

The Scalar Fields

The well-known equation for the scalar field is the Klein-Gordon (K-G) equation. The Klein-Gordon equation is a relativistic wave equation, obtained earlier for a single particle but it was initially unacceptable for the following two reasons. It admitted not only negative energy states but also negative probability densities. The latter was a more serious defect and was physically unacceptable. Only when Pauli and Weisskopf[1] found a way out and reinterpreted the Klein-Gordon equation as a field equation in the same sense as Maxwell's equations for electromagnetic field, the interest revived and became a subject of intense study.

Klein-Gordon equations are used to describe particles of spin 0 such as π mesons and K mesons. Real *i.e.* Hermitian fields represent uncharged particles and complex (non-Hermitian) fields represent charged particles.

7.1 One-component Real Field

First let us write down the K-G equation in different forms as are found in the literature.

$$(\Box + \kappa^2)\Phi(\mathbf{x}) = 0; \quad (\partial^\mu \partial_\mu + \kappa^2)\Phi(\mathbf{x}) = 0; \tag{7.1}$$

where

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2. \tag{7.2}$$

[1] W. Pauli and V. Weisskopf, *Helv. Phys. Acta.*, **7**, 709 (1934).

The symbols ∂^μ, ∂_μ and κ denote

$$\partial^\mu = \frac{\partial}{\partial x_\mu}; \quad \partial_\mu = \frac{\partial}{\partial x^\mu}; \qquad (\mu = 0, 1, 2, 3) \tag{7.3}$$

$$\kappa = \frac{mc}{\hbar}. \qquad \text{(Inverse Compton wavelength of the particle)} \tag{7.4}$$

Hereafter, we shall use natural units ($\hbar = c = 1$). In natural units, the K-G equation (7.1) can be rewritten as

$$(\partial^\mu \partial_\mu + m^2)\Phi(\mathbf{x}) = 0. \tag{7.5}$$

The plane wave solution of K-G Eq. $(\partial^\mu \partial_\mu + m^2)u_k(\mathbf{x}) = 0$ is

$$u_k(\mathbf{x}) = \frac{1}{\sqrt{V}} \frac{1}{\sqrt{2\omega_k}} e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{7.6}$$

where V is the quantization volume and

$$\mathbf{k}\cdot\mathbf{x} = k_0 x_0 - \boldsymbol{k}\cdot\boldsymbol{x} = \omega_k t - \boldsymbol{k}\cdot\boldsymbol{x}, \tag{7.7}$$

$$\mathbf{k}^2 = k^\mu k_\mu = \omega_k^2 - k^2 - = m^2, \tag{7.8}$$

$$\omega_k^2 = k^2 + m^2. \tag{7.9}$$

We use the convention of representing the four-vectors by bold up-right letters $(\mathbf{k}, \mathbf{x})$, three vectors by bold italics $(\boldsymbol{k}, \boldsymbol{x})$ and the scalars by ordinary italics (k, x).

7.1.1 Fourier decomposition of the field

The field operator $\Phi(\mathbf{x})$ in Eq. (7.5) is expanded in terms of the complete set of $u_k(\mathbf{x})$, defined by Eq. (7.6).

$$\begin{aligned} \Phi(\mathbf{x}) &= \sum_k \left(a_k u_k(\mathbf{x}) + a_k^\dagger u_k^*(\mathbf{x}) \right) \\ &= \frac{1}{\sqrt{V}} \sum_k \frac{1}{\sqrt{2\omega_k}} \left(a(\boldsymbol{k})\, e^{-i\mathbf{k}\cdot\mathbf{x}} + a^\dagger(\boldsymbol{k})\, e^{i\mathbf{k}\cdot\mathbf{x}} \right) \end{aligned} \tag{7.10}$$

The second term on the right hand side of Eq. (7.10) is the hermitian conjugate of the first term and is included in order to make the field function $\Phi(\mathbf{x})$ real. The expansion coefficients[2] a_k and $a_k^\dagger$ will have to be interpreted as the annihilation and creation operators and the field

[2]The expansion coefficients are denoted either by $a_k, a_k^\dagger$ or by $a(\boldsymbol{k}), a^\dagger(\boldsymbol{k})$.

operator $\Phi(\mathbf{x})$ represents an assembly of particles. More often, Eq. (7.10) is split into two terms

$$\Phi(\mathbf{x}) = \Phi^+(\mathbf{x}) + \Phi^-(\mathbf{x}), \tag{7.11}$$

where $\Phi^+(\mathbf{x})$ which carries the annihilation operator a_k corresponds to the positive energy part of the decomposition of the field function $\Phi(\mathbf{x})$ and $\Phi^-(\mathbf{x})$ which carries the creation operator $a_k^\dagger$ corresponds to the negative energy part of the decomposition of the field function $\Phi(\mathbf{x})$. Using the prescription

$$\frac{1}{\sqrt{V}} \sum_k \longrightarrow \frac{1}{(2\pi)^{3/2}} \int d^3k.$$

we can make a transition from the discrete values of $\boldsymbol{k}$ to continuous values and write down explicitly $\Phi^+(\mathbf{x})$ and $\Phi^-(\mathbf{x})$.

$$\Phi^+(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int_{k_0>0} \frac{d^3k}{\sqrt{2\omega_k}} a(\boldsymbol{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}; \tag{7.12}$$

$$\Phi^-(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int_{k_0>0} \frac{d^3k}{\sqrt{2\omega_k}} a^\dagger(\boldsymbol{k}) e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{7.13}$$

The splitting of the field function Φ into two parts $\Phi^+(\mathbf{x})$ and $\Phi^-(\mathbf{x})$ that correspond to positive and negative frequency parts, is introduced here, since we will have many occasions to use them at a later stage in our discussion.

Since we are using the Lagrangian formalism, the Lagrangian density will play a central role and so let us conjecture the Lagrangian density $\mathscr{L}_{KG}$ that will satisfy the Euler-Lagrange equation and yield the Klein-Gordon equation.

$$\begin{aligned} \mathscr{L}_{KG}(\mathbf{x}) &= \frac{1}{2} \left\{ \sum_\mu \left(\frac{\partial \Phi}{\partial x_\mu} \right)^2 - m^2 \Phi^2 \right\} \\ &= \frac{1}{2} \left(\partial^\mu \Phi(\mathbf{x})\, \partial_\mu \Phi(\mathbf{x}) - m^2 \Phi^2(\mathbf{x}) \right). \end{aligned} \tag{7.14}$$

The repeated index μ implies summation over μ. Given the above Lagrangian density, one can find the conjugate momentum density $\Pi(\mathbf{x})$ and the Hamiltonian density $\mathscr{H}(\mathbf{x})$ by the established procedure.

$$\Pi(\mathbf{x}) = \frac{\partial \mathscr{L}_{KG}}{\partial \dot{\Phi}} = \partial_0 \Phi = \dot{\Phi}. \tag{7.15}$$

$$\begin{aligned}\mathscr{H}(\mathbf{x}) &= \Pi(\mathbf{x})\dot{\Phi}(\mathbf{x}) - \mathscr{L}_{KG}(\mathbf{x}) \\ &= \partial^0\Phi\,\partial_0\Phi - \frac{1}{2}\left(\partial^\mu\Phi(\mathbf{x})\,\partial_\mu\Phi(\mathbf{x}) - m^2\Phi^2(\mathbf{x})\right) \\ &= \frac{1}{2}\left\{\dot{\Phi}^2 + (\boldsymbol{\nabla}\Phi)^2 + m^2\Phi^2\right\}. \qquad (7.16)\end{aligned}$$

In Eq. (7.16), $\boldsymbol{\nabla}$ denotes the ordinary three-dimensional gradient. The energy density as given by Eq. (7.16) is a positive-definite quantity as should be expected.

The total energy of the field is[3]

$$\begin{aligned}H &= \int \mathscr{H}\, d^3x = \int \left(\Pi\dot{\Phi} - \mathscr{L}_{KG}\right) d^3x \\ &= \int \left(\frac{\partial\mathscr{L}}{\partial\dot{\Phi}}\dot{\Phi} - \mathscr{L}\right) d^3x = \int \left(\frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\Phi}{\partial x_0}\right)}\frac{\partial\Phi}{\partial x_0} - \mathscr{L}\right) d^3x \\ &= \int T_{00}\, d^3x, \qquad (7.17)\end{aligned}$$

where T_{00} is the (0,0) component of the canonical energy momentum tensor $T_{\mu\nu}$ defined earlier.

$$T_{\mu\nu} = \frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\Phi}{\partial x_\nu}\right)}\frac{\partial\Phi}{\partial x_\mu} - \mathscr{L}\delta_{\mu\nu}. \qquad (7.18)$$

Using the Fourier decomposition (7.10) of the field, the conjugate momentum density $\Pi(\mathbf{x})$ given by Eq. (7.15) can be written in the expanded form using annihilation and creation operators $a(\boldsymbol{k})$ and $a^\dagger(\boldsymbol{k})$.

$$\Pi(\mathbf{x}) = \dot{\Phi}(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_k \frac{-i\omega_k}{\sqrt{2\omega_k}}\left(a(\boldsymbol{k})\, e^{-i\mathbf{k}\cdot\mathbf{x}} - a^\dagger(\boldsymbol{k})\, e^{i\mathbf{k}\cdot\mathbf{x}}\right). \qquad (7.19)$$

Since the commutation relations between Φ and Π and the Hamiltonian H do not depend on time, let us choose $t = 0$. Then,

$$\Phi(\boldsymbol{x}, 0) = \frac{1}{\sqrt{V}}\sum_k \frac{1}{\sqrt{2\omega_k}}\left(a(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} + a^\dagger(\boldsymbol{k})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right). \qquad (7.20)$$

[3]Since we are considering only the K-G field in his chapter, we will hereafter omit the suffix KG for the $\mathscr{L}$.

A transition from the discrete values of k to continuous values of k can be made by using the prescription (*vide* Solved Problem (7.4))

$$\frac{1}{\sqrt{V}}\sum_{k} \longrightarrow \frac{1}{(2\pi)^{3/2}}\int d^3k.$$

$$\Phi(\boldsymbol{x},0) = \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2\omega_k}}\left(a(\boldsymbol{k})\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}} + a^\dagger(\boldsymbol{k})\,e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right). \tag{7.21}$$

$$\Pi(\boldsymbol{x},0) = \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2\omega_k}}(-i\omega_k)\left(a(\boldsymbol{k})\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}} - a^\dagger(\boldsymbol{k})\,e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right). \tag{7.22}$$

Taking the Fourier inversions of (7.21) and (7.22), we obtain (*vide* Solved Problem (7.2))

$$a(\boldsymbol{k}) = \frac{1}{(2\pi)^{3/2}}\int \frac{d^3x}{\sqrt{2\omega_k}}\,e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\left\{\omega_k\Phi(\boldsymbol{x}) + i\Pi(\boldsymbol{x})\right\}. \tag{7.23}$$

$$a^\dagger(\boldsymbol{k}) = \frac{1}{(2\pi)^{3/2}}\int \frac{d^3x}{\sqrt{2\omega_k}}\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\left\{\omega_k\Phi(\boldsymbol{x}) - i\Pi(\boldsymbol{x})\right\}. \tag{7.24}$$

7.1.2 Quantization of the scalar field

The scalar field is quantized by invoking the equal-time commutation relations between the field function $\Phi(\mathbf{x})$ and its conjugate momentum function $\Pi(\mathbf{x})$.

$$[\Phi(\boldsymbol{x},t),\Pi(\boldsymbol{x}',t)]_- = i\hbar\,\delta(\boldsymbol{x}-\boldsymbol{x}') = i\delta(\boldsymbol{x}-\boldsymbol{x}'); \tag{7.25}$$

$$[\Phi(\boldsymbol{x},t),\Phi(\boldsymbol{x}',t)]_- = [\Pi(\boldsymbol{x},t),\Pi(\boldsymbol{x}',t)]_- = 0; \tag{7.26}$$

They, in turn, yield the commutation relations between the annihilation and creation operators. (*vide* Solved Problem (7.3))

$$[a(\boldsymbol{k}),a^\dagger(\boldsymbol{k}')]_- = \delta_{\boldsymbol{k},\boldsymbol{k}'}; \tag{7.27}$$

$$[a(\boldsymbol{k}),a(\boldsymbol{k}')]_- = [a^\dagger(\boldsymbol{k}),a^\dagger(\boldsymbol{k}')]_- = 0. \tag{7.28}$$

There is a very close analogy between the simple harmonic oscillator[4] and the field. The operators $a(\boldsymbol{k})$ and $a^\dagger(\boldsymbol{k})$ are interpreted as the annihilation and creation operators for a field quantum with momentum $\boldsymbol{k}$. The number operator $N(\boldsymbol{k})$ is defined as

$$N(\boldsymbol{k}) = a^\dagger(\boldsymbol{k})a(\boldsymbol{k}), \tag{7.29}$$

[4]The reader is referred to: V. Devanathan, *Quantum Mechanics*, Sec. 6.4.3, Narosa Publishing House, New Delhi (2005), for further details on simple harmonic oscillator.

which operating on the basic state in the Fock space

$$|n_1, n_2, n_3, \cdots, n_k, \cdots\rangle \tag{7.30}$$

yields the number of quanta associated with the momentum $\boldsymbol{k}$.

$$N(\boldsymbol{k})|n_1, n_2, n_3, \cdots, n_k, \cdots\rangle = n_k|n_1, n_2, n_3, \cdots, n_k, \cdots\rangle. \tag{7.31}$$

Using Eq. (7.16), it can be shown that the total Hamiltonian H is (*vide* Solved Problem (7.8))

$$\begin{aligned} H &= \int \mathscr{H} d^3x = \frac{1}{2}\int \left(\Pi^2 + (\boldsymbol{\nabla}\Phi)^2 + m^2\Phi^2\right) d^3x \\ &= \frac{1}{2}\sum_k \omega_k \left(a^\dagger(\boldsymbol{k})a(\boldsymbol{k}) + a(\boldsymbol{k})a^\dagger(\boldsymbol{k})\right) \\ &= \sum_k \omega_k \left(a^\dagger(\boldsymbol{k})a(\boldsymbol{k}) + \frac{1}{2}\right). \end{aligned} \tag{7.32}$$

In the last step, the commutation relation (7.27) is used. The ground state is one with $n_k = 0$ for all k and we should expect the ground state energy to be zero. But on the other hand, it becomes infinity $E = \sum_k \frac{1}{2}\omega_k$. The second quantized Klein-Gordon field thus gives a description of an assembly of infinitely many identical non-interacting particles of the same mass. They are not only spinless but also uncharged. That the associated quanta are uncharged follows from the fact that the four-vector current density corresponding to the Hermitian field is zero.

7.1.3 Ground state and normal ordering

The ground state of the field which is called the vacuum is defined as

$$a(\boldsymbol{k})|0\rangle = 0, \quad \text{for all } \boldsymbol{k}. \tag{7.33}$$

Assuming the normalization of the ground state

$$\langle 0|0\rangle = 1, \tag{7.34}$$

we find the expectation value of the total Hamiltonian in the ground state to be

$$\langle 0|H|0\rangle = \frac{1}{2}\sum_{k=0}^{\infty} \omega_k, \tag{7.35}$$

which is infinite. This is because, we have one oscillator for each value of momentum and we have an infinite number of such oscillators, each having a zero point energy. Thus the ground state energy becomes infinite. Now, let us redefine the zero of energy in such a way as to make the ground state energy zero. This can be done by a simple prescription called *normal ordering.* Whenever, we have a product of annihilation and creation operators, simply arrange all the annihilation operators to the right of all creation operators as if the commutators were zero. This is called normal ordering. Once we have done this arrangement, then the operators are treated once again as operators with the usual commutation relations. Using this prescription of normal ordering, we find the expectation value of the normal-ordered Hamiltonian denoted by the symbol $: H :$ or $N()$ to be

$$\langle\Psi| : H : |\Psi\rangle = \sum_{\boldsymbol{k}} \omega_k \langle\Psi|a^\dagger(\boldsymbol{k})a(\boldsymbol{k})|\Psi\rangle = \sum_{\boldsymbol{k}} \omega_k |a(\boldsymbol{k})|\Psi\rangle|^2. \quad (7.36)$$

The energy thus obtained is always positive and the expectation value of the energy of the vacuum now becomes

$$\langle 0| : H : |0\rangle = 0, \quad (7.37)$$

so that it corresponds to the ground state energy which is zero.

7.2 Complex Scalar Field

Any complex scalar field can be represented by two independent fields, one corresponding to the real part and the other corresponding to the imaginary part or alternatively by the complex field $\Phi(\mathbf{x})$ and its complex conjugate $\Phi^*(\mathbf{x})$, each of which obey the Klein-Gordon equation as given below:

$$\left(\partial^\mu \partial_\mu + m^2\right) \Phi(\mathbf{x}) = 0; \qquad \left(\partial^\mu \partial_\mu + m^2\right) \Phi^*(\mathbf{x}) = 0. \quad (7.38)$$

The Lagrangian density for such a complex field is given by

$$\mathscr{L}(\mathbf{x}) = \partial^\mu \Phi(\mathbf{x})\, \partial_\mu \Phi^*(\mathbf{x}) - m^2 \Phi(\mathbf{x}) \Phi^*(\mathbf{x}), \quad (7.39)$$

where $\Phi(\mathbf{x})$ and $\Phi^*(\mathbf{x})$ are to be treated as independent functions which can be treated as a linear combination of two independent real fields $\Phi_1(\mathbf{x})$ and $\Phi_2(\mathbf{x})$.

$$\Phi(\mathbf{x}) = \frac{1}{\sqrt{2}} \left(\Phi_1(\mathbf{x}) + i\Phi_2(\mathbf{x})\right); \quad \Phi^*(\mathbf{x}) = \frac{1}{\sqrt{2}} \left(\Phi_1(\mathbf{x}) - i\Phi_2(\mathbf{x})\right). \quad (7.40)$$

Each of these fields obey the Klein-Gordon equation.

$$\left(\partial^\mu \partial_\mu + m^2\right) \Phi_r(\mathbf{x}) = 0, \qquad r = 1, 2. \tag{7.41}$$

The Lagrangian density for the complex field can also be written as the sum of the Lagrangian densities of the two independent real fields Φ_1 and Φ_2.

$$\mathscr{L} = \frac{1}{2} \sum_{r=1}^{2} \left(\partial^\mu \Phi_r \, \partial_\mu \Phi_r - m^2 \Phi_r\right), \qquad r = 1, 2. \tag{7.42}$$

The conjugate momentum density of each field is given by

$$\Pi_r(\mathbf{x}) = \frac{\partial \mathscr{L}}{\partial \dot{\Phi}_r} = \partial_0 \Phi_r = \dot{\Phi}_r, \qquad r = 1, 2. \tag{7.43}$$

The Hamiltonian density of the complex field is given by

$$\mathscr{H} = \sum_{r=1}^{2} \Pi_r(\mathbf{x}) \dot{\Phi}_r(\mathbf{x}) - \mathscr{L}(\mathbf{x}), \tag{7.44}$$

from which the total Hamiltonian H is obtained by performing the three-dimensional space integration.

$$\begin{aligned} H &= \int \mathscr{H} d^3x \\ &= \sum_{r=1}^{2} \left\{ \int \partial^0 \Phi_r \, \partial_0 \Phi_r \, d^3x - \frac{1}{2} \int \left(\partial^\mu \Phi_r \, \partial_\mu \Phi_r - m^2\right) d^3x \right\} \end{aligned} \tag{7.45}$$

As was done earlier for the real scalar field, we can do the Fourier decomposition of the field for Φ_1 and Φ_2.

$$\Phi_r(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_k \frac{1}{\sqrt{2\omega_k}} \left\{ a_r(\boldsymbol{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} + a_r^\dagger(\boldsymbol{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \right\}, \tag{7.46}$$

where the operators $a_r(\boldsymbol{k})$ and $a_r^\dagger(\boldsymbol{k})$ are the annihilation and creation operators which obey the commutation relations

$$\left[a_r(\boldsymbol{k}), a_s^\dagger(\boldsymbol{k}')\right]_- = \delta_{rs} \, \delta(\boldsymbol{k} - \boldsymbol{k}'); \tag{7.47}$$

$$\left[a_r(\boldsymbol{k}), a_s(\boldsymbol{k}')\right]_- = 0 = \left[a_r^\dagger(\boldsymbol{k}), a_s^\dagger(\boldsymbol{k})\right]_- . \tag{7.48}$$

The commutators indicate that the field represents a system of bosons. The complex scalar field represents two types of particles and this would become clear by looking at the Fourier expansion of Φ using the relations (7.40).

$$\Phi(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_k \frac{1}{\sqrt{2\omega_k}} \left\{ a(\boldsymbol{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} + b^\dagger(\boldsymbol{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \right\}, \tag{7.49}$$

where

$$a(\boldsymbol{k}) = \frac{1}{\sqrt{2}} \left(a_1(\boldsymbol{k}) + i a_2(\boldsymbol{k}) \right), \qquad b(\boldsymbol{k}) = \frac{1}{\sqrt{2}} \left(a_1(\boldsymbol{k}) - i a_2(\boldsymbol{k}) \right). \tag{7.50}$$

The operators $a(\boldsymbol{k})$ and $b(\boldsymbol{k})$ and their Hermitian conjugates satisfy the commutation relations

$$\left[a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}') \right]_- = \left[b(\boldsymbol{k}), b^\dagger(\boldsymbol{k}') \right]_- = \delta(\boldsymbol{k} - \boldsymbol{k}'). \tag{7.51}$$

All other commutators with different combinations vanish.

From Eq. (7.49), we can at once write down the Fourier expansion for $\Phi^*(\mathbf{x})$.

$$\Phi^*(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_k \frac{1}{\sqrt{2\omega_k}} \left\{ a^\dagger(\boldsymbol{k}) e^{i\mathbf{k}\cdot\mathbf{x}} + b(\boldsymbol{k}) e^{-i\mathbf{k}\cdot\mathbf{x}} \right\}, \tag{7.52}$$

A complex scalar field is, in fact, consists of two fields $\Phi_1(\mathbf{x})$ and $\Phi_2(\mathbf{x})$ and $a_1^\dagger$ and $a_2^\dagger$ are the creation operators that produce the quanta that correspond to these fields. It is clear that a complex scalar field represents two different sets of particles. Instead of the fields $\Phi_1(\mathbf{x})$ and $\Phi_2(\mathbf{x})$, we can as well choose $\Phi(\mathbf{x})$ and $\Phi^*(\mathbf{x})$. For the latter choice, $a^\dagger$ and $b^\dagger$ are the the creation operators for the field quanta that are described by the fields $\Phi(\mathbf{x})$ and $\Phi^*(\mathbf{x})$. The latter choice is preferable since the Lagrangian (7.39) is invariant under the transformation

$$\Phi \to \Phi\, e^{ie\alpha}, \qquad \Phi^* \to \Phi^*\, e^{-ie\alpha}. \tag{7.53}$$

This transformation leads to the conservation of charge and current of the fields.

7.2.1 Charge-current density

Define the four vector current $j^\mu(\mathbf{x})$ by

$$\begin{aligned} j^\mu(\mathbf{x}) &= -ie \left(\frac{\partial \mathscr{L}}{\partial \Phi_{,\mu}} \Phi - \frac{\partial \mathscr{L}}{\partial \Phi^*_{,\mu}} \Phi^* \right) \\ &= ie \left\{ \left(\partial_\mu \Phi^* \right) \Phi - \left(\partial_\mu \Phi \right) \Phi^* \right\}. \end{aligned} \tag{7.54}$$

Using the Klein-Gordon field Eqs. (7.38), we get

$$\partial_\mu j^\mu(\mathbf{x}) = iem^2[\Phi^*(\mathbf{x}), \Phi(\mathbf{x})] = 0 \tag{7.55}$$

The four-vector $j^\mu(\mathbf{x})$ is to be interpreted as the charge-current density four vector and j^0 is the electric charge density.

The total charge of the field is

$$\begin{aligned} Q &= \int j^0(\mathbf{x}) d^3x \\ &= e\int \{(\partial_0\Phi^*)\Phi - (\partial_0\Phi)\Phi^*\}\, d^3x, \quad \text{using Eq. (7.54)} \\ &= \frac{e}{2}\sum_{kk'}\Big[\{a(\boldsymbol{k}'), a^\dagger(\boldsymbol{k})\} - \{b(\boldsymbol{k}), b(\boldsymbol{k}')\}\Big] \\ &= e\sum_k (N_+(\boldsymbol{k}) - N_-(\boldsymbol{k}))\,, \end{aligned} \tag{7.56}$$

where

$$N_+(\boldsymbol{k}) = a^\dagger(\boldsymbol{k})a(\boldsymbol{k}) \quad \text{and} \quad N_-(\boldsymbol{k}) = b^\dagger(\boldsymbol{k})b(\boldsymbol{k}).$$

Note that $Q = 0$ if anticommutation relations (Fermi-Dirac statistics) are used.

The total charge of the field is

$$Q = e\sum_k (n_+(\boldsymbol{k}) - n_-(\boldsymbol{k}))\,, \tag{7.57}$$

with

$$n_\pm(\boldsymbol{k}) = 0, 1, 2, \cdots, \infty.$$

Thus the operators $a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}), N_+(\boldsymbol{k})$ denote the annihilation, creation and number operator for particle of electric charge e and the operators $b(\boldsymbol{k}), b^\dagger(\boldsymbol{k}), N_-(\boldsymbol{k})$ denote the corresponding operators for a particle of charge $-e$. The total Hamiltonian is given by

$$\begin{aligned} H &= \frac{1}{2}\sum_k \omega_k \sum_{r=1}^{2}\Big\{a_r(\boldsymbol{k}), a_r^\dagger(\boldsymbol{k})\Big\} \\ &\equiv \frac{1}{2}\sum_k \omega_k \Big[\Big\{a(\boldsymbol{k}), a^\dagger(\boldsymbol{k})\Big\} + \Big\{b(\boldsymbol{k}), b^\dagger(\boldsymbol{k})\Big\}\Big] \\ &= \sum_k [N_+(\boldsymbol{k}) + N_-(\boldsymbol{k})]\,\omega_k + H_0, \end{aligned} \tag{7.58}$$

with

$$H_0 = \sum_k \omega_k I.$$

The total field energy of the K-G field is

$$E_{KG} = E_+ + E_-, \tag{7.59}$$

where

$$E_\pm = \sum_k \left(n_\pm(\boldsymbol{k}) + \frac{1}{2} \right) \omega_k.$$

The total field energy would be zero if anticommutation relations are used for annihilation and creation operators. This means that the Klein-Gordon equation can be consistently quantized only by using the commutation relations corresponding to Bose-Einstein statistics.

7.2.2 Particles and antiparticles

From Eq. (7.57), we find that the complex scalar field indicates that there are two types of particles, one with charge $+e$ and the other with charge $-e$. Apart from the sign of the charge, both the particles have identical mass and other properties. So, they have to be associated with the particle and its antiparticle. This is a fundamental feature of relativistic quantum field theory, which is fully supported by experiments.

As an example of particle-antiparticle pair, we can cite π^+ and π^- mesons. They can be described by a complex Klein-Gordon field. For a real field, the charge operator Q is identically zero and such a field can describe π^0 meson.

It is possible to extend the above consideration to other types of charges, for instance, hypercharge. The invariance of Lagrangian density $\mathscr{L}$ under phase transformation similar to (7.53) will allow the conservation of hypercharge. This will allow the occurrence of particles and antiparticles differing from each other in the sign of hypercharge. So, even electrically neutral particle can have a distinguishable antiparticle as in the case of neutral K^0 and $\overline{K}^0$. They have opposite hypercharge $Y = \pm 1$ and they can be represented by a complex Klein-Gordon field. Although electric charge is conserved in all types of interaction, hypercharge is conserved in strong and electromagnetic interactions but violated in weak interactions.

7.3 The Covariant Commutation Relations

In relativistic field theories, calculations can be performed in such a way that the relativistic invariance of the theory can be exhibited at every stage. This is made possible by the introduction of the Lorentz invariant commutation relations.

The scalar field function $\Phi(\mathbf{x})$ can be written in a Lorentz invariant way as

$$\begin{aligned}\Phi(\mathbf{x}) &= \frac{1}{(2\pi)^{3/2}}\int_{k_0>0}\frac{d^3k}{\sqrt{2\omega_k}}\left(a(\boldsymbol{k})e^{-i\mathbf{k}\cdot\mathbf{x}}+a^\dagger(\boldsymbol{k})e^{+i\mathbf{k}\cdot\mathbf{x}}\right)\\ &= \Phi^+(\mathbf{x})+\Phi^-(\mathbf{x}).\end{aligned}\tag{7.60}$$

The destruction operator $a(\boldsymbol{k})$ has positive frequency dependence and the creation operator $a^\dagger(\boldsymbol{k})$ has negative frequency dependence. This distinction is Lorentz invariant and allows an invariant decomposition of $\Phi(\mathbf{x})$ into positive and negative frequency parts.

$$\Phi^+(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int_{k_0>0}\frac{d^3k}{\sqrt{2\omega_k}}a(\boldsymbol{k})e^{-i\mathbf{k}\cdot\mathbf{x}},\tag{7.61}$$

$$\Phi^-(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int_{k_0>0}\frac{d^3k}{\sqrt{2\omega_k}}a^\dagger(\boldsymbol{k})e^{i\mathbf{k}\cdot\mathbf{x}}.\tag{7.62}$$

It can be verified that

$$\{\Phi^-(\mathbf{x})\}^\dagger=\Phi^+(\mathbf{x}).$$

$\Phi^+(\mathbf{x})$ is a destruction operator and $\Phi^-(\mathbf{x})$ is a creation operator. The vacuum state $|0\rangle$ which has been characterized by $a(\boldsymbol{k})|0\rangle=0$ for all $\boldsymbol{k}$ has now the property

$$\Phi^+(\mathbf{x})|0\rangle=0.$$

Let us now evaluate the commutator $[\Phi(\mathbf{x}),\Phi(\mathbf{x}')]_-$.

$$\begin{aligned}&[\Phi(\mathbf{x}),\Phi(\mathbf{x}')]_-\\ &= \Phi(\mathbf{x})\Phi(\mathbf{x}')-\Phi(\mathbf{x}')\Phi(\mathbf{x})\\ &= \frac{1}{(2\pi)^{3/2}}\int_{k_0>0}\frac{d^3k}{\sqrt{2\omega_k}}\int_{k_0'>0}\frac{d^3k'}{\sqrt{2\omega_{k'}}}\\ &\quad\times\left\{[a(\boldsymbol{k}),a^\dagger(\boldsymbol{k}')]_-e^{i\mathbf{k}'\cdot\mathbf{x}'}e^{-i\mathbf{k}\cdot\mathbf{x}}+[a^\dagger(\boldsymbol{k}),a(\boldsymbol{k}')]_-e^{-i\mathbf{k}'\cdot\mathbf{x}'}e^{i\mathbf{k}\cdot\mathbf{x}}\right\}\end{aligned}\tag{7.63}$$

Using the commutation relations between the annihilation and creation operators

$$\begin{aligned}[a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- &= \delta(\boldsymbol{k}-\boldsymbol{k}'),\\ [a(\boldsymbol{k}), a(\boldsymbol{k}')]_- &= [a^\dagger(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- = 0,\end{aligned}$$

and integrating over d^3k', we get after replacing ω_k by k_0 (since $\omega_k^2 = k_0^2 = k^2 + m^2$)

$$\begin{aligned}[\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_- &= \frac{1}{2(2\pi)^3}\int_{k_0>0}\frac{d^3k}{k_0}\left(e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} - e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\right)\\ &= i\Delta(\mathbf{x}-\mathbf{x}'),\end{aligned} \tag{7.64}$$

with

$$\Delta(\mathbf{x}-\mathbf{x}') = -\frac{i}{2(2\pi)^3}\int_{k_0>0}\frac{d^3k}{k_0}\left(e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} - e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\right). \tag{7.65}$$

Since

$$[\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_- = -[\Phi(\mathbf{x}'), \Phi(\mathbf{x})]_-,$$

we get

$$\Delta(\mathbf{x}-\mathbf{x}') = -\Delta(\mathbf{x}'-\mathbf{x}). \tag{7.66}$$

It can be shown that $\Delta(\mathbf{x}-\mathbf{x}')$ satisfies the Klein-Gordon equation.

$$(\Box_{\mathbf{x}} + \kappa^2)\Delta(\mathbf{x}-\mathbf{x}') = 0. \tag{7.67}$$

The two-point function $\Delta(\mathbf{x}-\mathbf{x}')$ can be considered as the propagator of the scalar field.

It can be shown that

$$[\Phi^+(\mathbf{x}), \Phi^+(\mathbf{x}')]_- = [\Phi^-(\mathbf{x}), \Phi^-(\mathbf{x}')]_- = 0. \tag{7.68}$$

In a similar way, we can also find the commutators $[\Phi^+(\mathbf{x}), \Phi^-(\mathbf{x}')]_-$ and $[\Phi^-(\mathbf{x}), \Phi^+(\mathbf{x}')]_-$.

$$\begin{aligned}&[\Phi^+(\mathbf{x}), \Phi^-(\mathbf{x})]_-\\ &= \Phi^+(\mathbf{x})\Phi^-(\mathbf{x}) - \Phi^-(\mathbf{x})\Phi^+(\mathbf{x})\\ &= \frac{1}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{\sqrt{2\omega_k}}\int_{k_0'>0}\frac{d^3k'}{\sqrt{2\omega_{k'}}}\\ &\quad\times\left\{a(\boldsymbol{k})a^\dagger(\boldsymbol{k}')e^{-i\mathbf{k}\cdot\mathbf{x}}e^{i\mathbf{k}'\cdot\mathbf{x}'} - a^\dagger(\boldsymbol{k}')a(\boldsymbol{k})e^{i\mathbf{k}'\cdot\mathbf{x}'}e^{-i\mathbf{k}\cdot\mathbf{x}}\right\}.\end{aligned} \tag{7.69}$$

Using the commutation relation

$$[a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- = \delta(\boldsymbol{k} - \boldsymbol{k}'),$$

and integrating over d^3k', we get after replacing ω_k by k_0

$$\begin{aligned}[\Phi^+(\mathbf{x}), \Phi^-(\mathbf{x})]_- &= \frac{1}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{2k_0}e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \\ &= i\Delta^+(\mathbf{x}-\mathbf{x}'). \end{aligned} \tag{7.70}$$

In a similar way, we get

$$\begin{aligned}[\Phi^-(\mathbf{x}), \Phi^+(\mathbf{x}')]_- &= -\frac{1}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{2k_0}e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} \\ &= i\Delta^-(\mathbf{x}-\mathbf{x}'). \end{aligned} \tag{7.71}$$

Thus, we find

$$\Delta^+(\mathbf{x}-\mathbf{x}') + \Delta^-(\mathbf{x}-\mathbf{x}') = \Delta(\mathbf{x}-\mathbf{x}'). \tag{7.72}$$

Likewise,

$$\begin{aligned}&\Delta^+(\mathbf{x}-\mathbf{x}') - \Delta^-(\mathbf{x}-\mathbf{x}') \\ &= -\frac{i}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{2k_0}\left\{e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} + e^{+i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\right\} \\ &= -\frac{i}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{2k_0}e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}\left\{e^{-ik_0(x_0-x_0')} + e^{+ik_0(x_0-x_0')}\right\} \\ &= -\frac{i}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{2k_0}e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}2\cos k_0(x_0-x_0') \\ &= -i\Delta^{(1)}(\mathbf{x}-\mathbf{x}'). \end{aligned} \tag{7.73}$$

From Eqs. (7.72) and (7.73), it follows that

$$\Delta^+(\mathbf{x}-\mathbf{x}') = \frac{1}{2}\left\{\Delta(\mathbf{x}-\mathbf{x}') - i\Delta^{(1)}(\mathbf{x}-\mathbf{x}')\right\}; \tag{7.74}$$

$$\Delta^-(\mathbf{x}-\mathbf{x}') = \frac{1}{2}\left\{\Delta(\mathbf{x}-\mathbf{x}') + i\Delta^{(1)}(\mathbf{x}-\mathbf{x}')\right\}; \tag{7.75}$$

The covariant commutation relations play a crucial role in treating physical problems and performing calculations in a relativistically invariant way.

Review Questions

7.1 Write dowm the Klein-Gordon equation for a real scalar field and discuss how it can be Fourier analyzed and quantized to show that it represents an assembly of Bose particles.

7.2 Show that the ground state of the system that the Klein-Gordon field represents is of infinite energy and explain how this difficulty can be overcome by normal ordering.

7.3 Show how the complex scalar field represents an assembly of charged particles whereas the real scalar field represents an assembly of uncharged particles.

7.4 Find the following covariant commutation relations between the scalar field functions at two different space-time points $\mathbf{x}$ and $\mathbf{x}'$:

$$[\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_-, \quad [\Phi^+(\mathbf{x}), \Phi^-(\mathbf{x}')]_-, \quad [\Phi^-(\mathbf{x}), \Phi^+(\mathbf{x}')]_-,$$

where Φ^+ and Φ^- denote the positive and negative parts of the field functions. Explain their physical significance.

Problems

7.1 Given the Lagrangian density

$$\mathscr{L}_{KG}(\mathbf{x}) = \frac{1}{2}\left\{\sum_\mu \left(\frac{\partial \Phi}{\partial x_\mu}\right)^2 - m^2\Phi^2\right\},$$

deduce the Klein-Gordon field equation.

7.2 Given the field function $\Phi(\boldsymbol{x})$ and the corresponding conjugate momentum function $\Pi(\boldsymbol{x})$ for the Klein-Gordon field

$$\begin{aligned}\Phi(\boldsymbol{x}) &= \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2\omega_k}}\left(a(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} + a^\dagger(\boldsymbol{k})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right);\\ \Pi(\boldsymbol{x}) &= \frac{1}{(2\pi)^{3/2}}\int \frac{d^3k}{\sqrt{2\omega_k}}(-i\omega_k)\left(a(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} - a^\dagger(\boldsymbol{k})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right);\end{aligned}$$

obtain, by Fourier inversion, expressions for annihilation and creation operators $a(\boldsymbol{k})$ and $a^\dagger(\boldsymbol{k})$.

7.3 Show that the equal-time commutation relations

$$\begin{aligned}[\Phi(\mathbf{x}), \Pi(\mathbf{x}')]_- &= i\hbar\,\delta(\mathbf{x}-\mathbf{x}') = i\delta(\mathbf{x}-\mathbf{x}');\\ [\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_- &= [\Pi(\mathbf{x}), \Pi(\mathbf{x}')]_- = 0;\end{aligned}$$

yield the commutation relations between the annihilation and creation operators.

$$\begin{aligned}[a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- &= \delta_{k,k'};\\ [a(\boldsymbol{k}), a(\boldsymbol{k}')]_- &= [a^\dagger(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- = 0.\end{aligned}$$

7.4 It is given that the prescription for going from discrete momentum to continuous momentum in the Fourier expansion of the field operator is

$$\frac{1}{\sqrt{V}}\sum_k \longrightarrow \frac{1}{(2\pi)^{3/2}}\int d^3k.$$

Deduce the above prescription.

7.5 Show that the covariant commutation relation given by Eq. (7.64) vanishes for equal times. Explain its physical significance.

7.6 Starting from the covariant commutation relation (7.64), deduce the canonical commutation relation for equal times:

$$[\Pi(\boldsymbol{x},t),\phi(\boldsymbol{x}',t)]_- = -i\delta(\boldsymbol{x}-\boldsymbol{x}').$$

7.7 Demonstrate that the field function $\Phi(\mathbf{x})$ given by Eq. (7.60) is really invariant under Lorentz transformation and the splitting of $\Phi(\mathbf{x})$ into positive and negative frequency parts $\Phi^+(\mathbf{x})$ and $\Phi^-(\mathbf{x})$ are also Lorentz invariant.

7.8 Given the Hamiltonian density $\mathscr{H}$ of the scalar field (Eq. (7.16),

$$\mathscr{H}(\boldsymbol{x}) = \frac{1}{2}\left(\Pi^2 + (\boldsymbol{\nabla}\Phi)^2 + m^2\Phi^2\right),$$

deduce the expression (7.32) for the total Hamiltonian H.

$$H = \int \mathscr{H}(\boldsymbol{x})d^3x = \frac{1}{2}\sum_k \omega_k \left\{a^\dagger(\boldsymbol{k})a(\boldsymbol{k}) + a(\boldsymbol{k})a^\dagger(\boldsymbol{k})\right\}.$$

Solutions to Problems

7.1 Given the Lagrangian density of the field, the field should obey the Euler-Lagrange equation

$$\frac{\partial\mathscr{L}}{\partial\Phi} - \frac{\partial}{\partial x^\mu}\frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\Phi}{\partial x^\mu}\right)} = 0. \tag{7.76}$$

The first term yields

$$\frac{\partial\mathscr{L}}{\partial\Phi} = -m^2\Phi. \tag{7.77}$$

The second term gives

$$\frac{\partial}{\partial x^\mu}\frac{\partial\mathscr{L}}{\partial\left(\frac{\partial\Phi}{\partial x^\mu}\right)} = \frac{\partial}{\partial x^\mu}\left\{\frac{\partial\Phi}{\partial x^\mu}\right\} = \frac{\partial^2\Phi}{\partial x^\mu\partial x^\mu}. \tag{7.78}$$

Substituting these in Euler-Lagrange equation, we get

$$\sum_{\mu} \frac{\partial^2 \Phi}{\partial x^\mu \partial x^\mu} + m^2 \Phi = 0 \tag{7.79}$$

This is the Klein-Gordon equation

$$(\Box + m^2)\Phi = 0, \tag{7.80}$$

with

$$\Box = \frac{\partial^2}{\partial x^{0\,2}} - \frac{\partial^2}{\partial x^{1\,2}} - \frac{\partial^2}{\partial x^{2\,2}} - \frac{\partial^2}{\partial x^{3\,2}} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2.$$

7.2 Multiply $\Phi(\boldsymbol{x})$ by ω_k and $\Pi(\boldsymbol{x})$ by i and then adding and subtracting the two expressions, we get

$$\omega_k \Phi(\boldsymbol{x}) + i\Pi(\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega_k}} 2\omega_k a(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{7.81}$$

$$\omega_k \Phi(\boldsymbol{x}) - i\Pi(\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3k}{\sqrt{2\omega_k}} 2\omega_k a^\dagger(\boldsymbol{k}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{7.82}$$

Taking the Fourier transforms of (7.81) and (7.82), we get

$$a(\boldsymbol{k}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3x}{\sqrt{2\omega_k}} e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \left\{\omega_k \Phi(\boldsymbol{x}) + i\Pi(\boldsymbol{x})\right\}. \tag{7.83}$$

$$a^\dagger(\boldsymbol{k}) = \frac{1}{(2\pi)^{3/2}} \int \frac{d^3x}{\sqrt{2\omega_k}} e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \left\{\omega_k \Phi(\boldsymbol{x}) - i\Pi(\boldsymbol{x})\right\}. \tag{7.84}$$

7.3 Using the expressions (7.83) and (7.84) for $a(\boldsymbol{k})$ and $a^\dagger(\boldsymbol{k}')$, we get

$$\begin{aligned} a(\boldsymbol{k})a^\dagger(\boldsymbol{k}') = {} & \frac{1}{(2\pi)^3} \int \frac{d^3x}{\sqrt{2\omega_k}} \frac{d^3x'}{\sqrt{2\omega_{k'}}} e^{i(\boldsymbol{k}'\cdot\boldsymbol{x}' - \boldsymbol{k}\cdot\boldsymbol{x})} \\ & \left\{\omega_k \Phi(\boldsymbol{x}) + i\Pi(\boldsymbol{x})\right\} \left\{\omega'_k \Phi(\boldsymbol{x}') - i\Pi(\boldsymbol{x}')\right\}. \end{aligned} \tag{7.85}$$

$$\begin{aligned} a^\dagger(\boldsymbol{k}')a(\boldsymbol{k}) = {} & \frac{1}{(2\pi)^3} \int \frac{d^3x}{\sqrt{2\omega_k}} \frac{d^3x'}{\sqrt{2\omega_{k'}}} e^{i(\boldsymbol{k}'\cdot\boldsymbol{x}' - \boldsymbol{k}\cdot\boldsymbol{x})} \\ & \left\{\omega_{k'} \Phi(\boldsymbol{x}') - i\Pi(\boldsymbol{x}')\right\} \left\{\omega_k \Phi(\boldsymbol{x}) + i\Pi(\boldsymbol{x})\right\}. \end{aligned} \tag{7.86}$$

The commutator

$$\begin{aligned} [a(\boldsymbol{k})a^\dagger(\boldsymbol{k}')]_- = {} & \frac{1}{(2\pi)^3} \int \frac{d^3x}{\sqrt{2\omega_k}} \frac{d^3x'}{\sqrt{2\omega_{k'}}} e^{i(\boldsymbol{k}'\cdot\boldsymbol{x}' - \boldsymbol{k}\cdot\boldsymbol{x})} \\ & \times \{\omega_k \omega_{k'} [\Phi(\boldsymbol{x}), \Phi(\boldsymbol{x}')]_- + [\Pi(\boldsymbol{x}), \Pi(\boldsymbol{x}')]_- \\ & -i\omega_k [\Phi(\boldsymbol{x}), \Pi(\boldsymbol{x}')]_- + i\omega_{k'} [\Pi(\boldsymbol{x}), \Phi(\boldsymbol{x}')]_-\}. \end{aligned} \tag{7.87}$$

Substituting the commutation relations between the field functions

$$[\Phi(\boldsymbol{x}), \Pi(\boldsymbol{x})]_- = i\delta(\boldsymbol{x} - \boldsymbol{x}'), \qquad [\Phi(\boldsymbol{x}), \Phi(\boldsymbol{x}')]_- = [\Pi(\boldsymbol{x}), \Pi(\boldsymbol{x}')]_- = 0,$$

The quantity within the curly bracket in Eq. (7.87) becomes

$$\{\cdots\} = \omega_k \delta(\boldsymbol{x} - \boldsymbol{x}') + \omega_{k'} \delta(\boldsymbol{x} - \boldsymbol{x}').$$

The integration over d^3x' is trivial and is done by replacing $\boldsymbol{x}'$ by $\boldsymbol{x}$ in the integrand.

$$\begin{aligned} [a(\boldsymbol{k})a^\dagger(\boldsymbol{k}')]_- &= \frac{1}{(2\pi)^3} \int \frac{d^3x}{\sqrt{2\omega_k}\sqrt{2\omega_{k'}}} e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{x}} \{\omega_k + \omega_{k'}\} \\ &= \delta(\boldsymbol{k} - \boldsymbol{k}'), \end{aligned} \tag{7.88}$$

since

$$\frac{1}{(2\pi)^3} \int d^3x e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{x}} = \delta(\boldsymbol{k} - \boldsymbol{k}').$$

7.4 The number of states available with momentum lying between k and $k+dk$ is

$$\frac{V d^3k}{h^3} = \frac{V d^3k}{(2\pi\hbar)^3},$$

where V is the normalization volume. The total number of states available in the case of momentum becoming a continuos variable is

$$\int \frac{V d^3k}{2\pi\hbar)^3}.$$

So, the summation over the discrete number of states goes over into an integral if the momentum becomes a continuous variable.

$$\begin{aligned} \sum_k &\longrightarrow \int \frac{V d^3k}{(2\pi\hbar)^3} \\ \frac{1}{V}\sum_k &\longrightarrow \int \frac{d^3k}{(2\pi\hbar)^3} \xrightarrow{\text{in natural units}} \int \frac{d^3k}{(2\pi)^3}. \end{aligned}$$

If the nrmalization volume is V, we usually take the normaliztion factor as $1/\sqrt{V}$. Hence

$$\frac{1}{\sqrt{V}} \sum_k \longrightarrow \frac{1}{(2\pi)^{3/2}} \int d^3k.$$

7.5 The commutator can be expanded as

$$\begin{aligned} [\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_- &= \Phi(\mathbf{x})\Phi(\mathbf{x}') - \Phi(\mathbf{x}')\Phi(\mathbf{x}) \\ &= \frac{1}{(2\pi)^{3/2}} \int_{k_0>0} \frac{d^3k}{\sqrt{2\omega_k}} \int_{k'_0>0} \frac{d^3k'}{\sqrt{2\omega_{k'}}} \\ &\quad \times \left\{ [a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- e^{i\mathbf{k}'\cdot\mathbf{x}'} e^{-i\mathbf{k}\cdot\mathbf{x}} + [a^\dagger(\boldsymbol{k}), a(\boldsymbol{k}')]_- e^{-i\mathbf{k}'\cdot\mathbf{x}'} e^{i\mathbf{k}\cdot\mathbf{x}} \right\}. \end{aligned}$$

Using the commutation relations between the annihilation and creation operators

$$\begin{aligned}[a(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- &= \delta(\boldsymbol{k}-\boldsymbol{k}'),\\ [a(\boldsymbol{k}), a(\boldsymbol{k}')]_- &= [a^\dagger(\boldsymbol{k}), a^\dagger(\boldsymbol{k}')]_- = 0,\end{aligned}$$

and integrating over d^3k', we get

$$\begin{aligned}[\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_- &= \frac{1}{2(2\pi)^3}\int_{k_0>0}\frac{d^3k}{k_0}\left(e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} - e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\right)\\ &= i\Delta(\mathbf{x}-\mathbf{x}'),\end{aligned}$$

with

$$\begin{aligned}\Delta(\mathbf{x}-\mathbf{x}') &= -\frac{i}{2(2\pi)^3}\int_{k_0>0}\frac{d^3k}{k_0}\left(e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} - e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\right)\\ &= \frac{1}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{k_0}e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}\sin k_0(x_0-x_0')\\ &= 0, \qquad \text{if } x_0 = x_0'.\end{aligned}$$

Hence the commutator $[\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_-$ vanishes for equal times.

A vanishing commutator allows one to simultaneously diagonalize the two operators and permits to make precise observation simultaneously of the two observables corresponding to the operators. So, the measurement of the Hermitian field operators $\Phi(\boldsymbol{x}, t)$ and $\Phi(\boldsymbol{x}, t)$ cannot interfere with one another, since no light signal can connect these two space-time points. It follows that the commutator vanishes for any space-like intervals *i.e* for space-time intervals $\boldsymbol{x}_1 - \boldsymbol{x}_2 > c(t_1 - t_2)$.

7.6 Let us start with the commutation relation of the scalar field operators.

$$[\Phi(\mathbf{x}), \Phi(\mathbf{x}')]_- = i\hbar\Delta(\mathbf{x}-\mathbf{x}').$$

$$\frac{\partial}{\partial t}[\Phi(\boldsymbol{x}, t), \Phi(\boldsymbol{x}', t')]_- = [\Pi(\boldsymbol{x}, t), \Phi(\boldsymbol{x}', t')]_-,$$

since $\frac{\partial}{\partial t}\Phi(\boldsymbol{x}', t') = 0$ and $\frac{\partial}{\partial t}\Phi(\boldsymbol{x}, t) = \Pi$.

Hence

$$[\Pi(\boldsymbol{x}, t), \Phi(\boldsymbol{x}', t')]_- = \frac{\partial}{\partial t}[\Phi(\boldsymbol{x}, t), \Phi(\boldsymbol{x}', t')]_- = i\hbar\frac{\partial}{\partial t}\Delta(\mathbf{x}-\mathbf{x}'),$$

where

$$\begin{aligned}\Delta(\mathbf{x}-\mathbf{x}') &= -\frac{i}{2}\frac{1}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{k_0}\left(e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} - (e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\right)\\ &= -\frac{i}{2}\frac{1}{(2\pi)^3}\int_{k_0>0}\frac{d^3k}{k_0}e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')}\{2i\sin(k_0(x_0-x_0'))\}\end{aligned}$$

The only factors that depend on time are $x_0 = ct$ and $x_0' = ct'$. So, differentiating with respect to t, we get

$$\frac{\partial}{\partial t}\sin(k_0(x_0 - x_0')) = k_0\cos(k_0(x_0 - x_0')) = k_0, \quad \text{if } x_0 = x_0'.$$

Substituting this, we get

$$\begin{aligned}\Delta(\mathbf{x} - \mathbf{x}') &= \frac{1}{(2\pi)^3}\int d^3k e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \\ &= \delta(\boldsymbol{x} - \boldsymbol{x}').\end{aligned}$$

Thus, we get back the canonical commutation rule for equal times.

$$[\Pi(\boldsymbol{x}, t), \Phi(\boldsymbol{x}', t)]_- = i\hbar\delta(\boldsymbol{x} - \boldsymbol{x}').$$

7.7 Expressions (7.60) - (7.62) for $\Phi(\mathbf{x}), \Phi^+(\mathbf{x}), \Phi^-(\mathbf{x})$, are not explicitly in Lorentz invariant forms since they involve three dimensional vectors but they can be cast in terms of four vectors using the following identity.

$$\begin{aligned}\Phi(\mathbf{x}) &= \frac{1}{(2\pi)^{3/2}}\int_{k_0>0}\frac{d^3k}{\sqrt{2\omega_k}}\left(a(\boldsymbol{k})e^{-i\mathbf{k}\cdot\mathbf{x}} + a^\dagger(\boldsymbol{k})e^{+i\mathbf{k}\cdot\mathbf{x}}\right) \\ &= \Phi^+(\mathbf{x}) + \Phi^-(\mathbf{x}).\end{aligned}$$

First let us show that

$$\begin{aligned}\frac{d^4k}{(2\pi)^4}2\pi\delta(\mathbf{k}^2 - m^2)\theta(k_0) &= \frac{d^4k}{(2\pi)^3}\delta(k_0^2 - \omega_k^2)\theta(k_0), \\ &\quad \text{since } \mathbf{k}^2 - m^2 = k_0^2 - \boldsymbol{k}^2 - m^2 = k_0^2 - \omega^2 \\ &= \frac{d^4k}{(2\pi)^3}\delta[(k_0 + \omega_k)(k_0 - \omega_k)]\theta(k_0) \\ &= \frac{d^4k}{(2\pi)^3}\frac{1}{2k_0}[\delta(k_0 + \omega_k) + \delta(k_0 - \omega_k)]\theta(k_0), \\ &\quad \text{since } \delta(x^2 - a^2) = \frac{1}{2a}\{\delta(x + a) + \delta(x - a)\} \\ &= \frac{d^3k}{(2\pi)^3}\frac{dk_0}{2k_0}\delta(k_0 - \omega_k) \\ &= \frac{d^3k}{(2\pi)^3}\frac{1}{2\omega_k}.\end{aligned}$$

Using the above identity and adjusting the normalization factor, we can display the above quantities in a Lorentz invariant form.

7.8 The Hamiltonian density $\mathscr{H}$ (7.16) consists of three terms. To find the total Hamiltonian, one has to perform integration over d^3x on each term.

In the absence of interaction, the Hamiltonian will be constant in time. So, we can choose $t = 0$.

Let us do the integration over the first term in some detail. Using Eq. (7.22), we get

$$\int \Pi^2(\boldsymbol{x})d^3x = \frac{1}{(2\pi)^3}\int \frac{d^3k d^3k' d^3x}{\sqrt{2\omega_k 2\omega_{k'}}}\,(-\omega_k\omega_{k'}) \times \left\{a(\boldsymbol{k})e^{i\boldsymbol{k}\cdot\boldsymbol{x}} - a^\dagger(\boldsymbol{k})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\right\}\left\{a(\boldsymbol{k}')e^{i\boldsymbol{k}'\cdot\boldsymbol{x}} - a^\dagger(\boldsymbol{k}')e^{-i\boldsymbol{k}'\cdot\boldsymbol{x}}\right\}.$$

First let us perform integration over d^3x. The $\boldsymbol{x}$- dependent terms are only within the curly brackets.

$$\begin{aligned}\int d^3x\{\cdots\}\{\cdots\} &= \int d^3x\left[a(\boldsymbol{k})a(\boldsymbol{k}')e^{i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{x}} - a(\boldsymbol{k})a^\dagger(\boldsymbol{k}')e^{i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{x}}\right.\\ &\quad \left. -a^\dagger(\boldsymbol{k})a(\boldsymbol{k}')e^{-i(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{x}} + a^\dagger(\boldsymbol{k})a^\dagger(\boldsymbol{k}')e^{-i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{x}}\right]\\ &= (2\pi)^3\left[a(\boldsymbol{k})a(\boldsymbol{k}')\delta(\boldsymbol{k}+\boldsymbol{k}') - a(\boldsymbol{k})a^\dagger(\boldsymbol{k}')\delta(\boldsymbol{k}-\boldsymbol{k}')\right.\\ &\quad \left. -a^\dagger(\boldsymbol{k})a(\boldsymbol{k}')\delta(\boldsymbol{k}-\boldsymbol{k}') + a^\dagger(\boldsymbol{k})a^\dagger(\boldsymbol{k}')\delta(\boldsymbol{k}+\boldsymbol{k}')\right],\end{aligned}$$

where we have used the standard integral

$$\int e^{i(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{x}}d^3x = (2\pi)^3\delta(\boldsymbol{k}+\boldsymbol{k}').$$

The integration over d^3k' is trivial because of the delta function. Thus, we get

$$\begin{aligned}\int \Pi^2(\boldsymbol{x})d^3x &= \int d^3k\left(\frac{\omega_k}{2}\right)\left\{a(\boldsymbol{k})a^\dagger(\boldsymbol{k}) + a^\dagger(\boldsymbol{k})a(\boldsymbol{k})\right.\\ &\quad \left. -a(\boldsymbol{k})a(-\boldsymbol{k}) - a^\dagger(\boldsymbol{k})a^\dagger(-\boldsymbol{k})\right\}. \qquad (7.89)\end{aligned}$$

In the same way, the integrals over the other two terms can be evaluated and we give below only the final results.

$$\begin{aligned}\int (\boldsymbol{\nabla}\Phi(\boldsymbol{x}))^2 d^3x &= \int d^3k\left(\frac{k^2}{2\omega_k}\right)\left\{a(\boldsymbol{k})a^\dagger(\boldsymbol{k}) + a^\dagger(\boldsymbol{k})a(\boldsymbol{k})\right.\\ &\quad \left. +a(\boldsymbol{k})a(-\boldsymbol{k}) + a^\dagger(\boldsymbol{k})a^\dagger(-\boldsymbol{k})\right\}. \qquad (7.90)\\ m^2\int (\Phi(\boldsymbol{x}))^2 d^3x &= \int d^3k\left(\frac{m^2}{2\omega_k}\right)\left\{a(\boldsymbol{k})a^\dagger(\boldsymbol{k}) + a^\dagger(\boldsymbol{k})a(\boldsymbol{k})\right.\\ &\quad \left. +a(\boldsymbol{k})a(-\boldsymbol{k}) + a^\dagger(\boldsymbol{k})a^\dagger(-\boldsymbol{k})\right\}. \qquad (7.91)\end{aligned}$$

Adding (7.89), (7.90) and (7.91) and observing that $\omega_k^2 = k^2 + m^2$, we get

$$H = \frac{1}{2}\int \omega_k[a^\dagger(\boldsymbol{k})a(\boldsymbol{k}) + a(\boldsymbol{k})a^\dagger(\boldsymbol{k})].$$

Chapter 8

The Dirac Field

In the last chapter, we have considered the scalar fields which, when quantized, using commutation relations yield an assembly of particles which obey Bose-Einstein statistics. In this chapter, let us consider the Dirac spin-$\frac{1}{2}$ field which is described by a four-component spinor ψ. Each component ψ_α of the four-component spinor should be treated as independent dynamical variable. The Dirac field, when quantized, yields an assembly of spin-$\frac{1}{2}$ particles known as Fermions. They obey the Fermi-Dirac statistics and satisfy the Pauli exclusion principle. This is achieved by replacing the commutation relations between the annihilation and creation operators by anti-commutation relations.

It is found that Bosons which have integral spins can be quantized only by invoking commutation relations and the Fermions which have half-integral spins can be quantized only by invoking anticommutation relations. The close relation between spin and statistics and their connection to commutation or anticommutation relations between annihilation and creation operators are the essential features of relativistic quantum field theory.

The reader is advised to read in advance Chap. 2, where the Dirac equation is treated as a single particle relativistic wave equation. This is a prerequisite for studying the Dirac equation as a field equation with a field function that can be expanded in terms of a complete set of eigenfunctions of the free particle Dirac equation. Anyhow, a brief summary of the relevant results are given for the benefit of the reader.

8.1 Plane Wave Solutions of the Dirac Equation

Let us consider the Dirac equation in Feynman's notation[1] for a particle of mass m (in natural units $\hbar = c = 1$).

$$(\not{p} - m)\psi(\boldsymbol{x}, t) = 0 \quad \text{or} \quad (i\gamma^\mu \partial_\mu - m)\,\psi(\mathbf{x}) = 0. \tag{8.1}$$

The symbol ∂_μ stands for the differential $\partial/\partial x^\mu$ and $\mathbf{x}$ denotes the four-vector $x_0 = ct, x_1, x_2, x_3$. The repeated index μ implies summation over the index μ. They are, in total, four solutions for Eq. (8.1). Two of them correspond to positive energy E_p, one with spin up and the other with spin down. The other two correspond to negative energy $-E_p$, one with spin up and the other with spin down. The four plane wave solutions of Eq. (8.1) are[2]

$$\varphi_r(\mathbf{x}) = \frac{1}{\sqrt{V}} \frac{1}{\sqrt{2E_p}} u_r(\boldsymbol{p}) e^{-i\mathbf{p}\cdot\mathbf{x}}, \qquad r = 1, 2; \tag{8.2}$$

$$\varphi_r(\mathbf{x}) = \frac{1}{\sqrt{V}} \frac{1}{\sqrt{2E_p}} u_r(\boldsymbol{p}) e^{i\mathbf{p}\cdot\mathbf{x}}, \qquad r = 3, 4. \tag{8.3}$$

The scalar product of four vectors $\mathbf{p}$ and $\mathbf{x}$ are defined by

$$\mathbf{p}\cdot\mathbf{x} = Et - \boldsymbol{p}\cdot\boldsymbol{x}; \qquad \mathbf{p}^2 = p^\mu p_\mu = E^2 - \boldsymbol{p}^2 = m^2; \tag{8.4}$$

We use the convention of denoting the four-vectors by upright bold letters $(\mathbf{p}, \mathbf{x})$, three-vectors by italic bold letters $(\boldsymbol{p}, \boldsymbol{x})$ and scalars by ordinary italics (p, x). In Eqs. (8.2) and (8.3), $u_r(\boldsymbol{p})$ denotes the spinors which are really four-component but expressed below in the two-component form for the sake of compactness.

$$u_r(\boldsymbol{p}) = (E_p + m)^{1/2} \begin{bmatrix} \chi_\pm \\ \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + m}\chi_\pm \end{bmatrix}, \qquad r = 1, 2 \tag{8.5}$$

$$u_r(\boldsymbol{p}) = (E_p + m)^{1/2} \begin{bmatrix} \dfrac{\boldsymbol{\sigma}\cdot\boldsymbol{p}}{E_p + m}\chi_\pm \\ \chi_\pm \end{bmatrix}, \qquad r = 3, 4. \tag{8.6}$$

---

[1] The Feynman notation $\not{p}$ stands for

$$\not{p} = \gamma^\mu p_\mu = \gamma^0 E - \boldsymbol{\gamma}\cdot\boldsymbol{p} = i\left(\gamma^0 \frac{\partial}{\partial x^0} + \gamma^1 \frac{\partial}{\partial x^1} + \gamma^2 \frac{\partial}{\partial x^2} + \gamma_3 \frac{\partial}{\partial x^3}\right)$$
$$= i\sum_\mu \gamma^\mu \frac{\partial}{\partial x^\mu} = i\gamma^\mu \partial_\mu.$$

[2] The suffix r denotes the different spin orientations.

In Eqs. (8.5) and (8.6), the symbols E_p, χ_+ and χ_- stand for

$$E_p = +\sqrt{p^2+m^2}; \qquad \chi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \qquad \chi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{8.7}$$

Different authors use different normalizations for the spinors. We choose the following normalization for the spinors $u_r(\boldsymbol{p})$, their Hermitian conjugates $u_r^\dagger(\boldsymbol{p})$ and their Hermitian adjoints $\bar{u}_r(\boldsymbol{p}) = u_r^\dagger(\boldsymbol{p})\gamma_0$, the normalization volume being V.

$$\begin{aligned}
u_r^\dagger(\boldsymbol{p})u_s(\boldsymbol{p}) &= 2E_p\delta_{rs}; && r,s = 1,2,3,4. && (8.8)\\
\bar{u}_r(\boldsymbol{p})u_s(\boldsymbol{p}) &= 2m\delta_{rs}; && r,s = 1,2. && (8.9)\\
\bar{u}_r(\boldsymbol{p})u_s(\boldsymbol{p}) &= -2m\delta_{rs}; && r,s = 3,4. && (8.10)\\
\bar{u}_r(\boldsymbol{p})u_s(\boldsymbol{p}) &= 0 && r = 1,2;\ s = 3,4. && (8.11)\\
\bar{u}_r(\boldsymbol{p})u_s(\boldsymbol{p}) &= 0 && r = 3,4;\ s = 1,2. && (8.12)
\end{aligned}$$

The spinors satisfy the following equations:

$$\begin{aligned}
(\gamma^\mu p_\mu - m)\, u_r(\boldsymbol{p}) &= 0, && r = 1,2; && (8.13)\\
(\gamma^\mu p_\mu + m)\, u_r(\boldsymbol{p}) &= 0, && r = 3,4; && (8.14)
\end{aligned}$$

which are obtained from Eq. (8.1) using the relation

$$p_\mu = -i\hbar\frac{\partial}{\partial x^\mu} \xrightarrow[\text{units}]{\text{In natural}} -i\frac{\partial}{\partial x^\mu} = -i\partial_\mu.$$

The spinor normalizations are such that

$$\int_V \varphi_r^\dagger(\mathbf{x})\varphi_s(\mathbf{x})d^3x = 2E_p\delta_{rs}, \qquad r,s = 1,2,3,4.$$

The solutions $\varphi_1(\mathbf{x})$ and $\varphi_2(\mathbf{x})$ correspond to positive energy ($E = E_p$) with momentum $\boldsymbol{p}$ whereas the solutions $\varphi_3(\mathbf{x})$ and $\varphi_4(\mathbf{x})$ correspond to negative energy ($E = -E_p$) with momentum $-\boldsymbol{p}$.

The Hermitian adjoint of the Dirac equation (8.1) is

$$\bar{\psi}(\mathbf{x})\left(i\gamma^\mu \overleftarrow{\partial_\mu} + m\right) = 0, \tag{8.15}$$

with $\bar{\psi}(\mathbf{x}) = \psi^\dagger(\mathbf{x})\gamma_0$. The left-arrow sign indicates that the operator acts on the function located on the left.

8.2 Lagrangian Density for the Dirac Field

It can be easily verified that both the Dirac equation (8.1) and its Hermitian adjoint (8.15) can be obtained from the Lagrangian density

$$\mathscr{L}_D(\mathbf{x}) = \bar{\psi}(\mathbf{x})\left(i\gamma^\mu\partial_\mu - m\right)\psi(\mathbf{x}), \tag{8.16}$$

by insisting that it satisfies the Euler-Lagrange equation.

If ψ, is a complex field, then we can treat both ψ and $\bar{\psi}$ as independent fields. The Euler-Lagrange equation for $\bar{\psi}$ gives

$$\partial_\mu\left(\frac{\partial\mathscr{L}}{\partial\left(\partial_\mu\bar{\psi}\right)}\right) = \frac{\partial\mathscr{L}}{\partial\bar{\psi}}. \tag{8.17}$$

Since the Lagrangian density $\mathscr{L}$ given by Eq. (8.16) does not contain any derivative of $\bar{\psi}$, the left hand side of Eq. (8.17) vanishes and Eq. (8.17) yields the Dirac equation (8.1).

$$\frac{\partial\mathscr{L}}{\partial\bar{\psi}} = \left(i\gamma^\mu\partial_\mu - m\right)\psi(\mathbf{x}) = 0.$$

In a similar way, we can consider the Euler-Lagrange equation for ψ and obtain the Hermitian conjugate of the Dirac equation (8.15).

In quantum field theory, the Lagrangian density should be Hermitian. In the case of scalar fields, we found the corresponding Lagrangian to be Hermitian. But the Dirac Lagrangian (8.16) is not Hermitian. The mass term is Hermitian whereas the Hermitian conjugate of the other term is given by

$$\left(i\bar{\psi}\gamma^\mu\partial_\mu\psi\right)^\dagger = -i(\partial_\mu\psi)^\dagger\gamma^{\mu\dagger}\gamma_0\psi = -i(\partial_\mu\bar{\psi})\gamma^\mu\psi, \tag{8.18}$$

using the hermiticity property of the gamma matrices

$$(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0. \tag{8.19}$$

Instead of (8.16), we can use the following Hermitian Lagrangian density:

$$\mathscr{L}' = \frac{i}{2}\left(\bar{\psi}\gamma^\mu\partial_\mu\psi - (\partial_\mu\bar{\psi})\gamma^\mu\psi\right) - m\bar{\psi}\psi. \tag{8.20}$$

However, it is not necessary since the two Lagrangians $\mathscr{L}$ and $\mathscr{L}'$ are equivalent since they differ by a term which can be expressed as a total divergence. The difference

$$\mathscr{L} - \mathscr{L}' = \frac{i}{2}\left(\bar{\psi}\gamma^\mu\partial_\mu\psi + (\partial_\mu\bar{\psi})\gamma^\mu\psi\right) = \partial_\mu\left(\frac{i}{2}\bar{\psi}\gamma^\mu\psi\right), \tag{8.21}$$

is a total divergence. It may be recalled that the two Lagrangians which differ by a total divergence would yield the same equations of motion based on the action principle. So, we can continue to use the same Lagrangian (8.16).

Now, let us proceed to find the conjugate momentum density and the Hamiltonian density of the field.

$$\Pi(\mathbf{x}) = \frac{\partial \mathscr{L}_D}{\partial \dot{\psi}} = \frac{\partial \mathscr{L}_D}{\partial \psi_{,0}} = i\bar{\psi}(\mathbf{x})\gamma_0 = i\psi^\dagger(\mathbf{x}). \tag{8.22}$$

$$\begin{aligned} \mathscr{H}(\mathbf{x}) &= \Pi(\mathbf{x})\dot{\psi}(\mathbf{x}) - \mathscr{L}_D(\mathbf{x}) \\ &= \Pi(\mathbf{x})\dot{\psi}(\mathbf{x}) = i\psi^\dagger(\mathbf{x})\dot{\psi}(\mathbf{x}) = i\psi^\dagger(\mathbf{x})\frac{\partial\psi(\mathbf{x})}{\partial t}. \end{aligned} \tag{8.23}$$

The last line in Eq. (8.23) is obtained using the result that the Lagrangian density $\mathscr{L}_D$ given by Eq. (8.16) vanishes if the field function satisfies the Dirac equation.

Our aim is to find the total Hamiltonian H.

$$H = \int \mathscr{H} d^3x = \int \psi^\dagger(\mathbf{x}) i\frac{\partial\psi(\mathbf{x})}{\partial t} d^3x. \tag{8.24}$$

8.3 Fourier Decomposition of the Dirac Field

Let us now replace the negative energy electron spinors $u_3(-\boldsymbol{p})$ and $u_4(-\boldsymbol{p})$ by positive energy positron spinors $v_1(\boldsymbol{p})$ and $v_2(\boldsymbol{p})$,

$$v_1(\boldsymbol{p}) \equiv u_3(-\boldsymbol{p}), \qquad v_2(\boldsymbol{p}) \equiv u_4(-\boldsymbol{p}), \tag{8.25}$$

and treat the Dirac field functions $\psi(\mathbf{x})$ and $\bar{\psi}(\mathbf{x})$ as field operators. The field operators $\psi(\mathbf{x})$ and $\bar{\psi}(\mathbf{x})$ are expanded in terms of $\varphi_r(\mathbf{x})$, defined by Eqs. (8.2) and (8.3).

$$\psi(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\boldsymbol{p}}\sqrt{\frac{1}{2E_p}}\sum_{r=1}^{2}\left[c_r(\boldsymbol{p})u_r(\boldsymbol{p})e^{-i\mathbf{p}\cdot\mathbf{x}} + d_r^\dagger(\boldsymbol{p})v_r(\boldsymbol{p})e^{i\mathbf{p}\cdot\mathbf{x}}\right]; \tag{8.26}$$

$$\bar{\psi}(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\boldsymbol{p}}\sqrt{\frac{1}{2E_p}}\sum_{r=1}^{2}\left[c_r^\dagger(\boldsymbol{p})\bar{u}_r(\boldsymbol{p})e^{i\mathbf{p}\cdot\mathbf{x}} + d_r(\boldsymbol{p})\bar{v}_r(\boldsymbol{p})e^{-i\mathbf{p}\cdot\mathbf{x}}\right]. \tag{8.27}$$

The expansion coefficients $c_r(\boldsymbol{p})$ and $c_r^\dagger(\boldsymbol{p})$ are to be interpreted as annihilation and creation operators of the Fermions and the operators $d_r(\boldsymbol{p})$ and $d_r^\dagger(\boldsymbol{p})$ are to be interpreted as annihilation and creation operators

of the anti-Fermions. For this, let us write down $\psi^\dagger$ and $i\frac{\partial\psi}{\partial t}$, using the expansion (8.26) for $\psi(\mathbf{x})$.

$$\psi^\dagger(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_p \sqrt{\frac{1}{2E_p}}\sum_{r=1}^{2}\left[c_r^\dagger(\boldsymbol{p})u_r^\dagger(\boldsymbol{p})e^{i\mathbf{p}\cdot\mathbf{x}} + d_r(\boldsymbol{p})v_r^\dagger(\boldsymbol{p})e^{-i\mathbf{p}\cdot\mathbf{x}}\right]; \quad (8.28)$$

$$i\frac{\partial\psi(\mathbf{x})}{\partial t} = \frac{1}{\sqrt{V}}\sum_{p'}\sqrt{\frac{E_{p'}}{2}}\sum_{s=1}^{2}\Big[c_s(\boldsymbol{p}')u_s(\boldsymbol{p}')e^{-i\mathbf{p}'\cdot\mathbf{x}} - d_s^\dagger(\boldsymbol{p}')v_s(\boldsymbol{p}'))e^{i\mathbf{p}'\cdot\mathbf{x}}\Big]. \quad (8.29)$$

Substituting expressions (8.28) and (8.29) in Eq. (8.24), we get

$$\begin{aligned} H &= \frac{1}{V}\sum_{pp'}\sqrt{\frac{E_{p'}}{4E_p}}\int\sum_{r,s}\Big[c_r^\dagger(\boldsymbol{p})c_s(\boldsymbol{p}')u_r^\dagger(\boldsymbol{p})u_s(\boldsymbol{p}')e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} \\ &\quad - d_r(\boldsymbol{p})d_s^\dagger(\boldsymbol{p}')v_r^\dagger(\boldsymbol{p})v_s(\boldsymbol{p}')e^{-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} \\ &\quad - c_r^\dagger(\boldsymbol{p})d_s^\dagger(\boldsymbol{p}')u_r^\dagger(\boldsymbol{p})v_s(\boldsymbol{p}')e^{i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}} \\ &\quad + d_r(\boldsymbol{p})c_s(\boldsymbol{p}')v_r^\dagger(\boldsymbol{p})u_s(\boldsymbol{p}')e^{-i(\mathbf{p}+\mathbf{p}')\cdot\mathbf{x}}\Big]d^3x \quad (8.30) \\ &= \sum_p\sum_{r=1}^{2}E_p\left[c_r^\dagger(\boldsymbol{p})c_r(\boldsymbol{p}) - d_r(\boldsymbol{p})d_r^\dagger(\boldsymbol{p})\right]. \quad (8.31) \end{aligned}$$

For the reduction of (8.30) into (8.31), the reader may refer to the solved problem 8.1.

When $t = 0$, the exponential factors in Eqn. (8.30) involve only three-momenta $\boldsymbol{p}$ instead of four-momenta $\mathbf{p}$ and so the exponential factors reduce to

$$e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} \xrightarrow{t=0} e^{-i(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{x}}.$$

We can also replace the discrete summation over three-momenta by an integral in the case of continuous variation of three-momenta, using the following prescription.

$$\frac{1}{V}\sum_p \longrightarrow \int\frac{1}{(2\pi)^3}d^3p.$$

Let us now find the total charge that resides in the field. For this, let us first find the four-vector current $j^\mu(\mathbf{x})$ of the Dirac field.

$$\begin{aligned} j^\mu(\mathbf{x}) &= -ie\left(\frac{\partial\mathscr{L}}{\partial\psi_{,\mu}}\psi - \frac{\partial\mathscr{L}}{\partial\psi_{,\mu}^\dagger}\psi^\dagger\right) \\ &= e\bar{\psi}(\mathbf{x})\gamma^\mu\psi(\mathbf{x}). \quad (8.32) \end{aligned}$$

The zeroth component of the four-current j^μ yields the charge.

$$\begin{aligned} Q &= \int_V j^0(\mathbf{x}) d^3x \\ &= e\int_V \bar{\psi}(\mathbf{x})\gamma^0\psi(\mathbf{x})d^3x = e\int_V \psi^\dagger(\mathbf{x})\psi(\mathbf{x})d^3x \\ &= e\sum_{\boldsymbol{p}}\sum_{r=1}^{2}\left(c_r^\dagger(\boldsymbol{p})c_r(\boldsymbol{p}) + d_r(\boldsymbol{p})d_r^\dagger(\boldsymbol{p})\right). \end{aligned} \tag{8.33}$$

Expressions (8.31) and (8.33) give the total energy and the total charge of the Dirac field in terms of the annihilation and creation operators. Hitherto, no commutation relation has been assumed between the creation and annihilation operators. So, let us choose either the commutation or anticommutation relation between the annihilation and creation operators so that we get physically meaningful results.

The total energy will not be positive-definite if annihilation and creation operators commute. On the other hand, the choice of anticommutation relation between annihilation and creation operator makes the energy positive definite, although the zero point energy becomes negative. This should not cause any serious problem since we can redefine the energy scale with a convenient choice of zero point energy.

8.4 Quantization of the Dirac Field

The choice of commutation relation makes the total charge positive definite. On the other hand, the choice of anticommutation relation makes the total charge either positive or negative. This is physically acceptable. The particles and antiparticles have opposite charges and the net charge of the field is their difference.

Let us now write down the anticommutation relations between the annihilation and creation operators.

$$\begin{aligned} &\{c_r(\boldsymbol{p}), c_s^\dagger(\boldsymbol{p}')\}_+ = \delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}'); && \{d_r(\boldsymbol{p}), d_s^\dagger(\boldsymbol{p}')\}_+ = \delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}'); \\ &\{c_r(\boldsymbol{p}), c_s(\boldsymbol{p}')\}_+ = 0; && \{c_r^\dagger(\boldsymbol{p}), c_s^\dagger(\boldsymbol{p}')\}_+ = 0; \\ &\{d_r(\boldsymbol{p}), d_s(\boldsymbol{p}')\}_+ = 0; && \{d_r^\dagger(\boldsymbol{p}), c_s^\dagger(\boldsymbol{p}')\}_+ = 0. \end{aligned} \tag{8.34}$$

Using the anticommutation relations (8.34), the total energy and charge

of the field become

$$H = \sum_p \sum_{r=1}^{2} E_p \left(c_r^\dagger(\boldsymbol{p}) c_r(\boldsymbol{p}) + d_r^\dagger(\boldsymbol{p}) d_r(\boldsymbol{p}) \right) - \sum_p \sum_{r=1}^{2} E_p \quad (8.35)$$

$$= \sum_p \sum_{r=1}^{2} E_p \left(N_r(\boldsymbol{p}) + \tilde{N}_r(\boldsymbol{p}) \right) + E_0; \quad (8.36)$$

$$Q = e \sum_p \sum_{r=1}^{2} \left(N_r(\boldsymbol{p}) - \tilde{N}_r(\boldsymbol{p}) \right) + Q_0; \quad (8.37)$$

where N_r and $\tilde{N}_r$ denote the number operator for the particles and antiparticles respectively.

$$N_r(\boldsymbol{p}) = c_r^\dagger(\boldsymbol{p}) c_r(\boldsymbol{p}) \qquad \text{and} \qquad \tilde{N}_r(\boldsymbol{p}) = d_r^\dagger(\boldsymbol{p}) d_r(\boldsymbol{p}). \quad (8.38)$$

The zero-point energy and charge are denoted by E_0 and Q_0.

$$E_0 = -\sum_p \sum_{r=1}^{2} E_p = -2 \sum_p E_p; \qquad Q_0 = \sum_p \sum_{r=1}^{2} e = 2e \sum_p 1. \quad (8.39)$$

If we identify the zero-point energy and charge with the ground state or the vacuum state, then what is physically observable is the energy of a system with respect to the ground state. So, the physically measurable energy and charge of a system are given by

$$E = \sum_p \sum_{r=1}^{2} E_p \left(n_r(\boldsymbol{p}) + \tilde{n}_r(\boldsymbol{p}) \right); \quad (8.40)$$

$$Q = e \sum_p \sum_{r=1}^{2} \left(n_r(\boldsymbol{p}) - \tilde{n}_r(\boldsymbol{p}) \right); \quad (8.41)$$

where $n_r(\boldsymbol{p})$ and $\tilde{n}_r(\boldsymbol{p})$ denote the eigenvalues of the number operators N_r and $\tilde{N}_r$. The anticommutation relation restricts the eigenvalues of the number operator $n_r(\boldsymbol{p}), \tilde{n}_r(\boldsymbol{p}) = 0, 1$.

The zero point energy will be avoided if we resort to the normal ordering. The normal ordering for Fermion operators is to arrange the annihilation operators to the right of creation operators, assuming the anticommutators to be zero. In the case of Bosons, the normal ordering was done by assuming the commutators to be zero.

8.4.1 Anticommutator between Dirac field functions

Having established the anticommutation relations (8.34) between the annihilation and creation operators, the anticommutation relation between the Dirac field operators at equal times $\psi_\alpha(\boldsymbol{x})$ and $\bar{\psi}_\beta(\boldsymbol{x})$ can be deduced. Treating $\boldsymbol{p}$ as a continuous variable and putting $t = 0$, we get the following expressions from Eqs. (8.26) and (8.27).

$$\begin{aligned}
&\psi(\boldsymbol{x}, t=0) \\
&\quad = \frac{1}{(2\pi)^{3/2}} \int d^3p \frac{1}{2E_p} \sum_{r=1}^{2} \left\{ c_r(\boldsymbol{p}) u_r(\boldsymbol{p}) e^{i\boldsymbol{p}\cdot\boldsymbol{x}} + d_r^\dagger(\boldsymbol{p}) v_r(\boldsymbol{p}) e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} \right\}; \quad (8.42) \\
&\bar{\psi}(\boldsymbol{x}, t=0) \\
&\quad = \frac{1}{(2\pi)^{3/2}} \int d^3p \frac{1}{2E_p} \sum_{r=1}^{2} \left\{ c_r^\dagger(\boldsymbol{p}) \bar{u}_r(\boldsymbol{p}) e^{-i\boldsymbol{p}\cdot\boldsymbol{x}} + d_r(\boldsymbol{p}) \bar{v}_r(\boldsymbol{p}) e^{i\boldsymbol{p}\cdot\boldsymbol{x}} \right\}. \quad (8.43)
\end{aligned}$$

Since ψ is a four-component spinor, each component has to be treated as independent field variable[3]. Let us now evaluate the anticommutator $\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x})\}_+$.

$$\begin{aligned}
&\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x}')\}_+ \\
&\quad = \frac{1}{2\pi)^3} \int \frac{d^3p d^3p'}{\sqrt{2E_p 2E_{p'}}} \sum_{r,s} \Big\{ \{c_r(\boldsymbol{p}), c_s^\dagger(\boldsymbol{p}')\}_+ u_r^\alpha(\boldsymbol{p}) \bar{u}_s^\beta(\boldsymbol{p}') e^{i(\boldsymbol{p}\cdot\boldsymbol{x} - \boldsymbol{p}'\cdot\boldsymbol{x}')} \\
&\qquad + \{d_r^\alpha(\boldsymbol{p}), d_s^\beta(\boldsymbol{p}')\}_+ v_r^\alpha(\boldsymbol{p}) \bar{v}_s^\beta(\boldsymbol{p}') e^{-i(\boldsymbol{p}\cdot\boldsymbol{x} - \boldsymbol{p}'\cdot\boldsymbol{x}')} \Big\}; \\
&\quad = \frac{1}{2\pi)^3} \int \frac{d^3p}{2E_p} \sum_r \left\{ u_r^\alpha(\boldsymbol{p}) \bar{u}_r^\beta(\boldsymbol{p}) e^{i\boldsymbol{p}\cdot(\boldsymbol{x} - \boldsymbol{x}')} + v_r^\alpha(\boldsymbol{p}) \bar{v}_r^\beta(\boldsymbol{p}') e^{-i\boldsymbol{p}\cdot(\boldsymbol{x} - \boldsymbol{x}')} \right\}.
\end{aligned} \quad (8.44)$$

The last step is obtained using the anticommutation relations (8.34) for the annihilation and creation operators. The spinor indices α, β are used either as subscript or superscipt depending upon convenience. Since

$$\sum_r u_r^\alpha(\boldsymbol{p}) \bar{u}_r^\beta(\boldsymbol{p}) = (\not{p} + m)_{\alpha,\beta}, \qquad \sum_r v_r^\alpha(\boldsymbol{p}) \bar{v}_r^\beta(\boldsymbol{p}) = (\not{p} - m)_{\alpha,\beta},$$

[3]Since $\psi\bar{\psi}$ is a matrix and $\bar{\psi}\psi$ is a number, there is no meaning in trying to find the anticommutation relation $\{\psi(\boldsymbol{x}), \bar{\psi}(\boldsymbol{x})\}_+$. We can only find the anticommutation relation between their components $\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x})\}_+$.

we can rewrite Eq. (8.44) as

$$\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x}')\}_+ = \frac{1}{2\pi)^3} \int \frac{d^3p}{2E_p} \left\{ (\not{p} + m)_{\alpha\beta} e^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')} + (\not{p} - m)_{\alpha\beta} e^{-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \right\}. \tag{8.45}$$

Expanding

$$\not{p} + m = \gamma_0 E_p - \boldsymbol{\gamma}\cdot\boldsymbol{p} + m, \qquad \not{p} - m = \gamma_0 E_p - \boldsymbol{\gamma}\cdot\boldsymbol{p} - m,$$

and integrating over d^3p, we find that only the γ_0 term contributes, whereas the other terms cancel away[4]. Thus we get the anticommutation relations

$$\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x}')\}_+ = (\gamma_0)_{\alpha\beta}\,\delta(\boldsymbol{x}-\boldsymbol{x}'); \tag{8.46}$$

$$\{\psi_\alpha(\boldsymbol{x}), \psi^\dagger_\beta(\boldsymbol{x}')\}_+ = I_{\alpha,\beta}\delta(\boldsymbol{x}-\boldsymbol{x}'). \tag{8.47}$$

Since $\Pi(\boldsymbol{x}) = i\psi^\dagger(\boldsymbol{x})$ according to Eq. (8.22), we get the anticommutation relation

$$\{\psi_\alpha(\boldsymbol{x}), \Pi_\beta(\boldsymbol{x}')\}_+ = iI_{\alpha,\beta}\delta(\boldsymbol{x}-\boldsymbol{x}'), \tag{8.48}$$

where I is the unit matrix.

8.5 Covariant Anticommutation Relations

Just as we have obtained the covariant commutation relations for the scalar field operators in chapter 7, we can deduce the covariant anticommutation relations for the Dirac operators.

Let us start with the covariant form of the Dirac field operators $\psi(\mathbf{x})$ and $\bar{\psi}(\mathbf{x})$ and their Lorentz invariant divisions into positive frequency or energy parts $\psi^+(\mathbf{x}), \bar{\psi}^+(\mathbf{x})$ and negative frequency or energy parts $\psi^-(\mathbf{x}), \bar{\psi}^-(\mathbf{x})$.

$$\psi(\mathbf{x}) = \psi^+(\mathbf{x}) + \psi^-(\mathbf{x}); \qquad \bar{\psi}(\mathbf{x}) = \bar{\psi}^+(\mathbf{x}) + \bar{\psi}^-(\mathbf{x}); \tag{8.49}$$

[4]The following integral relations will be useful:

$$\int \cdot\,\boldsymbol{p}\, e^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}')} d^3p = \int \cdot\,\boldsymbol{p}'\, e^{i\boldsymbol{p}'\cdot(\boldsymbol{x}-\boldsymbol{x}')} d^3p' \text{ with } \boldsymbol{p}' = -\boldsymbol{p}.$$

$$\frac{1}{(2\pi)^3}\int e^{+i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}'} d^3p = \frac{1}{(2\pi)^3}\int e^{-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x}'} d^3p = \delta(\boldsymbol{x}-\boldsymbol{x}').$$

with

$$\psi^+(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int_{p_0>0}\frac{d^3p}{\sqrt{2E_p}}\sum_{r=1}^{2}c_r(\boldsymbol{p})u_r(\boldsymbol{p})\,e^{-i\mathbf{p}\cdot\mathbf{x}}; \quad (8.50)$$

$$\psi^-(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int_{p_0>0}\frac{d^3p}{\sqrt{2E_p}}\sum_{r=1}^{2}d_r^\dagger(\boldsymbol{p})v_r(\boldsymbol{p})\,e^{i\mathbf{p}\cdot\mathbf{x}}; \quad (8.51)$$

$$\bar\psi^+(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int_{p_0>0}\frac{d^3p}{\sqrt{2E_p}}\sum_{r=1}^{2}d_r(\boldsymbol{p})\bar v_r(\boldsymbol{p})\,e^{-i\mathbf{p}\cdot\mathbf{x}}; \quad (8.52)$$

$$\bar\psi^-(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}}\int_{p_0>0}\frac{d^3p}{\sqrt{2E_p}}\sum_{r=1}^{2}c_r^\dagger(\boldsymbol{p})\bar u_r(\boldsymbol{p})\,e^{i\mathbf{p}\cdot\mathbf{x}}. \quad (8.53)$$

These operators have some physical significance. $\psi^+(\mathbf{x})$ is the destruction operator for the Fermion, $\psi^-(\mathbf{x})$ is the creation operator for the anti-Fermion, $\bar\psi^+(\mathbf{x})$ is the destruction operator for the anti-Fermion and $\bar\psi^-(\mathbf{x})$ is the creation operator for the Fermion. The physical vacuum is characterized by the property that the destruction operators operating on the vacuum yield zero.

$$\psi^+(\mathbf{x})|0\rangle = 0; \qquad \bar\psi^+(\mathbf{x})|0\rangle = 0. \quad (8.54)$$

The anticommutation rules for the creation and destruction operators are

$$\{\psi^+(\mathbf{x}),\psi^-(\mathbf{x}')\}_+ = \{\bar\psi^-(\mathbf{x}),\bar\psi^+(\mathbf{x}')\}_+ = 0; \quad (8.55)$$

$$\{\bar\psi^+(\mathbf{x}),\psi^+(\mathbf{x}')\}_+ = \{\bar\psi^-(\mathbf{x}),\psi^-(\mathbf{x}')\}_+ = 0. \quad (8.56)$$

For the anticommutator $\{\psi_\alpha^+(\mathbf{x}),\bar\psi_\beta^-(\mathbf{x})\}_+$, we have

$$\{\psi_\alpha^+(\mathbf{x}),\bar\psi_\beta^-(\mathbf{x}')\}_+ = \frac{1}{(2\pi)^3}\int\frac{d^3p}{\sqrt{2E_p}}\int\frac{d^3p'}{\sqrt{2E_{p'}}}\sum_{r=1}^{2}\sum_{s=1}^{2}\{c_r(\boldsymbol{p}),c_s^\dagger(\boldsymbol{p}')\}_+ \times u_r^\alpha(\boldsymbol{p})\bar u_s^\beta(\boldsymbol{p}')e^{-i\mathbf{p}\cdot\mathbf{x}}e^{i\mathbf{p}'\cdot\mathbf{x}'}. \quad (8.57)$$

Since

$$\{c_r(\boldsymbol{p}),c_s^\dagger(\boldsymbol{p}')\}_+ = \delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}')$$

$$\sum_r u_r(\boldsymbol{p})\bar u_s(\boldsymbol{p}) = \not{p}+m,$$

we get

$$\begin{aligned}
\{\psi^+_\alpha(\mathbf{x}), \bar{\psi}^-_\beta(\mathbf{x}')\}_+ &= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E_p}\sum_{r=1}^{2} u^\alpha_r(\boldsymbol{p})\bar{u}^\beta_r(\boldsymbol{p})e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= \frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}(\not{p}+m)_{\alpha\beta}e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= (i\not{\nabla}_{\mathbf{x}}+m)_{\alpha\beta}\frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}e^{-i\mathbf{p}(\mathbf{x}-\mathbf{x}')} \\
&= i(i\not{\nabla}_{\mathbf{x}}+m)_{\alpha\beta}\Delta^+(\mathbf{x}-\mathbf{x}') \\
&= -iS^+_{\alpha\beta}(\mathbf{x}-\mathbf{x}'), \qquad (8.58)
\end{aligned}$$

where

$$\Delta^+(\mathbf{x}-\mathbf{x}') = \frac{-i}{2(2\pi)^3}\int_{p_0>0}\frac{d^3p}{p_0}e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \qquad (8.59)$$

$$S^+_{\alpha\beta}(\mathbf{x}-\mathbf{x}') = -(i\not{\nabla}_{\mathbf{x}}+m)_{\alpha\beta}\Delta^+(\mathbf{x}-\mathbf{x}'). \qquad (8.60)$$

In a similar way, we can deduce the anticommutation relation for the antiparticle operators.

$$\begin{aligned}
\{\psi^-_\alpha(\mathbf{x}), \bar{\psi}^+_\beta(\mathbf{x}')\}_+ &= \frac{1}{(2\pi)^3}\int \frac{d^3p}{\sqrt{2E_p}}\int \frac{d^3p'}{\sqrt{2E_{p'}}}\sum_{r=1}^{2}\sum_{s=1}^{2}\{d^\dagger_r(\boldsymbol{p}), d_s(\boldsymbol{p}')\}_+ \\
&\quad \times v^\alpha_r(\boldsymbol{p})\bar{v}^\beta_s(\boldsymbol{p}')e^{i\mathbf{p}\cdot\mathbf{x}}e^{-i\mathbf{p}'\cdot\mathbf{x}'}. \qquad (8.61)
\end{aligned}$$

Since

$$\begin{aligned}
\{d^\dagger_r(\boldsymbol{p}), d_s(\boldsymbol{p}')\}_+ &= \delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}'), \\
\sum_r v_r(\boldsymbol{p})\bar{v}_s(\boldsymbol{p}) &= \not{p}-m,
\end{aligned}$$

we get

$$\begin{aligned}
\{\psi^-_\alpha(\mathbf{x}), \bar{\psi}^+_\beta(\mathbf{x}')\}_+ &= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E_p}\sum_{r=1}^{2} v^\alpha_r(\boldsymbol{p})\bar{v}^\beta_r(\boldsymbol{p})e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= \frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}(\not{p}-m)_{\alpha\beta}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= (i\not{\nabla}_{\mathbf{x}}-m)_{\alpha\beta}\frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}e^{i\mathbf{p}(\mathbf{x}-\mathbf{x}')} \\
&= -i(i\not{\nabla}_{\mathbf{x}}-m)_{\alpha\beta}\Delta^-(\mathbf{x}-\mathbf{x}') \\
&= -iS^-_{\alpha\beta}(\mathbf{x}-\mathbf{x}'), \qquad (8.62)
\end{aligned}$$

where

$$\Delta^-(\mathbf{x}-\mathbf{x}') = \frac{i}{2(2\pi)^3}\int_{p_0>0}\frac{d^3p}{p_0}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \tag{8.63}$$

$$S^-_{\alpha\beta}(\mathbf{x}-\mathbf{x}') = (i\not{\nabla}_{\mathbf{x}}-m)_{\alpha\beta}\Delta^-(\mathbf{x}-\mathbf{x}'). \tag{8.64}$$

Finally, let us find the anticommutation relation between $\psi_\alpha(\mathbf{x})$ and $\bar{\psi}_\beta(\mathbf{x}')$.

$$\begin{aligned}\{\psi_\alpha(\mathbf{x}),\bar{\psi}_\beta(\mathbf{x}')\}_+ &= \left\{\left(\psi^+_\alpha(\mathbf{x})+\psi^-_\alpha(\mathbf{x})\right),\left(\bar{\psi}^+_\beta(\mathbf{x}')+\bar{\psi}^-_\beta(\mathbf{x}')\right)\right\}_+\\ &= \{\psi^+_\alpha(\mathbf{x}),\bar{\psi}^+_\beta(\mathbf{x}')\}_+ + \{\psi^+_\alpha(\mathbf{x}),\bar{\psi}^-_\beta(\mathbf{x}')\}_+\\ &\quad +\{\psi^-_\alpha(\mathbf{x}),\bar{\psi}^+_\beta(\mathbf{x}')\}_+ + \{\psi^-_\alpha(\mathbf{x}),\bar{\psi}^-_\beta(\mathbf{x}')\}_+ \end{aligned}\tag{8.65}$$

Two of the above anticommutator brackets vanish.

$$\{\psi^+_\alpha(\mathbf{x}),\bar{\psi}^+_\beta(\mathbf{x}')\}_+ = \{\psi^-_\alpha(\mathbf{x}),\bar{\psi}^-_\beta(\mathbf{x}')\}_+ = 0.$$

Using the results (8.58) and (8.62) for the other two, we get

$$\{\psi_\alpha(\mathbf{x}),\bar{\psi}_\beta(\mathbf{x}')\}_+ = -i[S^+_{\alpha\beta}(\mathbf{x}-\mathbf{x}')+S^-_{\alpha\beta}(\mathbf{x}-\mathbf{x}')]. \tag{8.66}$$

Review Questions

8.1 Write down the Lagrangian density for the Dirac field and obtain the field equations. Deduce expressions for the momentum density and the Hamiltonian density for the Dirac field.

8.2 Write down the Dirac field functions in terms of a complete set of single particle Dirac wave functions and arrive at an expression for the energy of an assembly of Dirac particles.

8.3 Deduce expressions for energy and charge of an assembly of Dirac particles in terms of creation and annihilation operators and show they yield meaningful results only when anticommutation relations are assumed for the annihilation and creation operators.

8.4 Given the anticommutation relations for the annihilation and creation operators for the Dirac particle

$$\{c_r(\boldsymbol{p}),c^\dagger_s(\boldsymbol{p}')\}_+ = \delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}'),\qquad \{d_r(\boldsymbol{p}),d^\dagger_s(\boldsymbol{p}')\}_+ = \delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}'),$$

deduce the following anticommutation relations for the Dirac field operators:

$$\begin{aligned}\{\psi_\alpha(\boldsymbol{x}),\bar{\psi}_\beta(\boldsymbol{x}')\}_+ &= \gamma^0_{\alpha,\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\\ \{\psi_\alpha(\boldsymbol{x}),\Pi_\beta(\boldsymbol{x}')\}_+ &= = iI_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x}')\end{aligned}$$

where I is the unit matrix.

8.5 Deduce the covariant anticommutation relations between the Dirac field operators:

$$\{\psi_\alpha^+(\mathbf{x}), \bar{\psi}_\beta^-(\mathbf{y})\}_+, \qquad \{\psi_\alpha^-(\mathbf{x}), \bar{\psi}_\beta^+(\mathbf{y})\}_+,$$

where the upper sign + and − indicates the positive and negative frequency part of the Dirac field operators.

Problems

8.1 Deduce the expression (8.31) for the energy of an assembly of Dirac particles from Eq. (8.30).

8.2 Deduce the expression (8.33) for the charge of the Dirac field in terms of the creation and annihilation operators.

8.3 Given the covariant anticommutation relations (8.58) and (8.62),

$$\begin{aligned}
\{\psi_\alpha^+(\mathbf{x}), \bar{\psi}_\beta^-(\mathbf{x}')\}_+ &= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E_p}\sum_{r=1}^{2} u_r^\alpha(\boldsymbol{p})\bar{u}_r^\beta(\boldsymbol{p})e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= \frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}(\not{p}+m)_{\alpha\beta}e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \\
\{\psi_\alpha^-(\mathbf{x}), \bar{\psi}_\beta^+(\mathbf{x}')\}_+ &= \frac{1}{(2\pi)^3}\int \frac{d^3p}{2E_p}\sum_{r=1}^{2} v_r^\alpha(\boldsymbol{p})\bar{v}_r^\beta(\boldsymbol{p})e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \\
&= \frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}(\not{p}-m)_{\alpha\beta}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}
\end{aligned}$$

deduce the anticommutation relation (8.46)

$$\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x}')\}_+ = (\gamma_0)_{\alpha\beta}\,\delta(\boldsymbol{x}-\boldsymbol{x}'),$$

for equal times or for $t - 0$

Solutions to Problems

8.1 Evaluate expression (8.30) at time $t = 0$. Then the expression involves only three-momentum $\boldsymbol{p}$ and $\boldsymbol{p}'$.

$$e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} \xrightarrow{t=0} e^{-i(\boldsymbol{p}-\boldsymbol{p}')\cdot\boldsymbol{x}}.$$

Using the orthonormal properties (8.8) - (8.12) of Dirac spinors, we get

$$\begin{aligned}
u_r^\dagger(\boldsymbol{p})u_s(\boldsymbol{p}') &= 2E_p\,\delta_{rs}\,\delta(\boldsymbol{p}-\boldsymbol{p}'); \\
v_r^\dagger(\boldsymbol{p})v_s(\boldsymbol{p}') &= 2E_p\,\delta_{rs}\,\delta(\boldsymbol{p}-\boldsymbol{p}'); \\
u_r^\dagger(\boldsymbol{p})v_s(\boldsymbol{p}') &= v_r^\dagger(\boldsymbol{p})u_s(\boldsymbol{p}') = 0.
\end{aligned}$$

Substituting them in expression (8.30) is equivalent to putting $r = s$ and $\boldsymbol{p} = \boldsymbol{p}'$.

$$\begin{aligned} H &= \frac{1}{V}\sum_{p}\sum_{r=1}^{2} E_p \left[c_r^\dagger(\boldsymbol{p})c_r(\boldsymbol{p}) - d_r(\boldsymbol{p})d_r^\dagger(\boldsymbol{p})\right]\int d^3x \\ &= \sum_{p}\sum_{r=1}^{2} E_p \left[c_r^\dagger(\boldsymbol{p})c_r(\boldsymbol{p}) - d_r(\boldsymbol{p})d_r^\dagger(\boldsymbol{p})\right]. \end{aligned}$$

8.2 The charge of the Dirac field is given by

$$Q = e\int_V \psi^\dagger(\mathbf{x})\psi(\mathbf{x})d^3x,$$

where $\psi^\dagger(\mathbf{x})\psi(\mathbf{x})$ is taken at the same space-time point ψ. The Fourier expansions of $\psi^\dagger(\mathbf{x})$ and $\psi(\mathbf{x})$ are given by Eqs. (8.28) and (8.26). They are reproduced here for the sake of convenience.

$$\begin{aligned} \psi^\dagger(\mathbf{x}) &= \frac{1}{\sqrt{V}}\sum_{p}\sqrt{\frac{1}{2E_p}}\sum_{r=1}^{2}\left[c_r^\dagger(\boldsymbol{p})u_r^\dagger(\boldsymbol{p})e^{i\mathbf{p}\cdot\mathbf{x}} + d_r(\boldsymbol{p})v_r^\dagger(\boldsymbol{p})e^{-i\mathbf{p}\cdot\mathbf{x}}\right]; \\ \psi(\mathbf{x}) &= \frac{1}{\sqrt{V}}\sum_{p'}\sqrt{\frac{1}{2E_{p'}}}\sum_{s=1}^{2}\left[c_s(\boldsymbol{p}')u_s(\boldsymbol{p}')e^{-i\mathbf{p}'\cdot\mathbf{x}} + d_s^\dagger(\boldsymbol{p}')v_s(\boldsymbol{p}')e^{i\mathbf{p}'\cdot\mathbf{x}}\right]; \end{aligned}$$

In taking the product of $\psi^\dagger(\mathbf{x})\psi(\mathbf{x})$, the cross terms involving the spinors $u_r^\dagger(\boldsymbol{p})v_s(\boldsymbol{p}')$ will vanish since they are orthogonal. Retaining only the contributing terms, we get

$$\begin{aligned} \psi^\dagger(\mathbf{x})\psi(\mathbf{x}) &= \frac{1}{V}\sum_{\boldsymbol{p},\boldsymbol{p}'}\frac{1}{2E_p 2E_{p'}}\sum_{r,s}\Big[c_r^\dagger(\boldsymbol{p})c_s(\boldsymbol{p}')u_r^\dagger(\boldsymbol{p})u_s(\boldsymbol{p}')e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}} \\ &\quad + d_r(\boldsymbol{p})d_s^\dagger(\boldsymbol{p}')v_r^\dagger(\boldsymbol{p})v_s(\boldsymbol{p}')e^{-i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}}\Big] \end{aligned}$$

Since

$$u_r^\dagger(\boldsymbol{p})u_s(\boldsymbol{p}') = 2E_p\delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}') \quad \text{and} \quad v_r^\dagger(\boldsymbol{p})v_s(\boldsymbol{p}') = 2E_p\delta_{rs}\delta(\boldsymbol{p}-\boldsymbol{p}')$$

we get

$$\psi^\dagger(\mathbf{x})\psi(\mathbf{x}) = \frac{1}{V}\sum_{r}[c_r^\dagger(\boldsymbol{p})c_s(\boldsymbol{p}) + d_r(\boldsymbol{p})d_r^\dagger(vp')]$$

Since $\int d^3x = V$ cancels with $1/V$, we get the simplified expression for Q.

$$Q = e\sum_{\boldsymbol{p}}\sum_{r}[c_r^\dagger(\boldsymbol{p})c_s(\boldsymbol{p}) + d_r(\boldsymbol{p})d_r^\dagger(vp').$$

8.3 The anticommutator $\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x'})\}_+$ can be expanded in term of positive and negative frequency operators.

$$\begin{aligned}\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x'})\}_+ &= \Big\{(\psi^+_\alpha(\boldsymbol{x}) + \psi^-_\alpha(\boldsymbol{x})), (\bar{\psi}^+_\beta(\boldsymbol{x'}) + \bar{\psi}^-_\beta(\boldsymbol{x'}))\Big\}_+ \\ &= \{\psi^+_\alpha(\boldsymbol{x}), \bar{\psi}^+_\beta(\boldsymbol{x'})\}_+ + \{\psi^+_\alpha(\boldsymbol{x}), \bar{\psi}^-_\beta(\boldsymbol{x'})\}_+ \\ &\quad + \{\psi^-_\alpha(\boldsymbol{x}), \bar{\psi}^+_\beta(\boldsymbol{x'})\}_+ + \{\psi^-_\alpha(\boldsymbol{x}), \bar{\psi}^-_\beta(\boldsymbol{x'})\}_+ \\ &= \{\psi^+_\alpha(\boldsymbol{x}), \bar{\psi}^-_\beta(\boldsymbol{x'})\}_+ + \{\psi^-_\alpha(\boldsymbol{x}), \bar{\psi}^+_\beta(\boldsymbol{x'})\}_+,\end{aligned}$$

since

$$\{\psi^+_\alpha(\boldsymbol{x}), \bar{\psi}^+_\beta(\boldsymbol{x'})\}_+ = \{\psi^-_\alpha(\boldsymbol{x}), \bar{\psi}^-_\beta(\boldsymbol{x'})\}_+ = 0.$$

It is given that

$$\begin{aligned}\{\psi^+_\alpha(\mathbf{x}), \bar{\psi}^-_\beta(\mathbf{x'})\}_+ &= \frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}(\not{p} + m)_{\alpha\beta} e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x'})}; \\ \{\psi^-_\alpha(\mathbf{x}), \bar{\psi}^+_\beta(\mathbf{x'})\}_+ &= \frac{1}{2(2\pi)^3}\int \frac{d^3p}{E_p}(\not{p} - m)_{\alpha\beta} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x'})}.\end{aligned}$$

For $t = 0$,

$$e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x'})} \to e^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x'})}; \qquad e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x'})} \to e^{-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x'})}$$

Please note that

$$\not{p} = \gamma_0 E - \boldsymbol{\gamma}\cdot\boldsymbol{p}.$$

Using the following integral identities

$$\begin{aligned}\frac{1}{(2\pi)^3}\int d^3p\, e^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x'})} &= \frac{1}{(2\pi)^3}\int d^3p\, e^{-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x'})} = \delta(\boldsymbol{x}-\boldsymbol{x'}), \\ \frac{1}{(2\pi)^3}\int d^3p(\gamma\cdot\boldsymbol{p}) e^{i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x'})} &= \frac{1}{(2\pi)^3}\int d^3p(-\gamma\cdot\boldsymbol{p}) e^{-i\boldsymbol{p}\cdot(\boldsymbol{x}-\boldsymbol{x'})},\end{aligned}$$

we find that the terms involving $\boldsymbol{\gamma}\cdot\boldsymbol{p}$ and m mutually cancel away and only the term involving γ_0 contributes. Thus we finally get

$$\begin{aligned}\{\psi_\alpha(\boldsymbol{x}), \bar{\psi}_\beta(\boldsymbol{x'})\}_+ &= \{\psi^+_\alpha(\boldsymbol{x}), \bar{\psi}^-_\beta(\boldsymbol{x'})\}_+ + \{\psi^-_\alpha(\boldsymbol{x}), \bar{\psi}^+_\beta(\boldsymbol{x'})\}_+, \\ &= (\gamma_0)_{\alpha\beta}\delta(\boldsymbol{x}-\boldsymbol{x'}).\end{aligned}$$

For simplicity, we have chosen $t = 0$ but this relation holds for any equal times $t = t'$.

Chapter 9

The Electromagnetic Field

The electromagnetic field is described by a set of Maxwell's equations which specify the electrical and magnetic field strengths $\boldsymbol{E}$ and $\boldsymbol{B}$ at each space point and their variation with time. Alternatively, it can be described in terms of scalar and vector potentials ϕ and $\boldsymbol{A}$. Since ϕ and $\boldsymbol{A}$ together yield 4 components, it is possible to represent them as a four-vector A^μ. The latter approach yields a field equation for a particle of zero rest mass similar to the familiar Klein-Gordon (K-G) equation. Earlier we have studied the K-G equation for a scalar field that corresponds to spin-zero particle. Here we consider a K-G equation for zero rest mass but with a four-vector function that corresponds to spin-1 particle.

$$\Box A^\mu(\mathbf{x}) = 0, \qquad \mu = 0, 1, 2, 3.$$

This is the relativistic field equation for the photon in free space.

The electrical and magnetic field strengths $\boldsymbol{E}$ and $\boldsymbol{B}$ together form 6 components and it is shown that they correspond to six independent components of an antisymmetric tensor $F^{\mu\nu}$ of rank 2 formed with A^μ. The scalar and vector potentials $\phi(\boldsymbol{x}, t)$ and $A(\boldsymbol{x}, t)$ are not observable quantities but the electrical and magnetic field strengths $\boldsymbol{E}(\boldsymbol{x}, t)$ and $\boldsymbol{B}(\boldsymbol{x}, t)$ that are derived from the potentials are measurable quantities. However it is found that the scalar and vector potentials are not unique but there exists more than one set of potentials that yield the same $\boldsymbol{E}$ and $\boldsymbol{B}$. This is often referred to as the invariance of the electromagnetic field under gauge transformation.

9.1 Maxwell's Equations

First let us write down Maxwell's equations for the electromagnetic field in unrationalized Gaussian (CGS) units, using the two different notations that are in vogue for differential operators in vector algebra.

$$\text{div}\,\boldsymbol{B} = 0 \qquad \text{or} \qquad \boldsymbol{\nabla}\cdot\boldsymbol{B} = 0. \tag{9.1}$$

$$\text{curl}\,\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0 \qquad \text{or} \qquad \boldsymbol{\nabla}\times\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t} = 0. \tag{9.2}$$

$$\text{curl}\,\boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{4\pi \boldsymbol{j}}{c} \qquad \text{or} \qquad \boldsymbol{\nabla}\times\boldsymbol{B} - \frac{1}{c}\frac{\partial \boldsymbol{E}}{\partial t} = \frac{4\pi \boldsymbol{j}}{c}. \tag{9.3}$$

$$\text{div}\,\boldsymbol{E} = 4\pi\rho \qquad \text{or} \qquad \boldsymbol{\nabla}\cdot\boldsymbol{E} = 4\pi\rho. \tag{9.4}$$

From Eq. (9.1), it can be inferred that $\boldsymbol{B}$ can be expressed as the curl of a vector $\boldsymbol{A}$ since div curl $\boldsymbol{A} = 0$, according to the vector algebra. Therefore

$$\boldsymbol{B} = \text{curl}\,\boldsymbol{A} = \boldsymbol{\nabla}\times\boldsymbol{A}, \tag{9.5}$$

where $\boldsymbol{A}$ is known as the vector potential of the electromagnetic field. Expressing $\boldsymbol{B}$ in terms of $\boldsymbol{A}$ and using the commutativity of the differential operators $\boldsymbol{\nabla}$ and $\frac{\partial}{\partial t}$, Eq. (9.2) can be rewritten as

$$\boldsymbol{\nabla}\times\left(\boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right) = 0. \tag{9.6}$$

Since curl grad $\phi = 0$, the quantity within bracket in Eq. (9.6) can be written as the gradient of a scalar function ϕ.

$$-\boldsymbol{\nabla}\phi = \boldsymbol{E} + \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}, \tag{9.7}$$

choosing the negative sign for convenience. Thus

$$\boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}. \tag{9.8}$$

By this procedure, we have expressed the fields $\boldsymbol{B}$ and $\boldsymbol{E}$ in terms of $\boldsymbol{A}$ and ϕ, known as the vector and the scalar potential describing the electromagnetic field.

Now let us express Eqs. (9.3) and (9.4) in terms of $\boldsymbol{A}$ and ϕ. Eq. (9.3) can be rewritten as (using Eqs. (9.5) and (9.8)).

$$\text{curl}\,\text{curl}\,\boldsymbol{A} - \frac{1}{c}\frac{\partial}{\partial t}\left(-\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right) = \frac{4\pi \boldsymbol{j}}{c}. \tag{9.9}$$

Using the identity

$$\text{curl}\,\text{curl}\,\boldsymbol{A} = \text{grad}\,\text{div}\,\boldsymbol{A} - \boldsymbol{\nabla}^2\boldsymbol{A}, \tag{9.10}$$

in vector algebra, Eq. (9.3) becomes

$$\begin{aligned} \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{A}) - \boldsymbol{\nabla}^2\boldsymbol{A} - \frac{1}{c}\frac{\partial}{\partial t}\left(-\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}\right) &= \frac{4\pi \boldsymbol{j}}{c} \\ \boldsymbol{\nabla}\left(\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\phi}{\partial t}\right) - \boldsymbol{\nabla}^2\boldsymbol{A} + \frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} &= \frac{4\pi\boldsymbol{j}}{c}. \end{aligned} \tag{9.11}$$

In the above Equation, put

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} + \frac{1}{c}\frac{\partial\phi}{\partial t} = 0. \tag{9.12}$$

Then Eq. (9.11) simplifies to

$$\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2} - \boldsymbol{\nabla}^2\boldsymbol{A} = \frac{4\pi\boldsymbol{j}}{c}. \tag{9.13}$$

Equation (9.13) is obtained under the condition (9.12), which is known as *the Lorentz condition.*

Substituting Eq. (9.8) for $\boldsymbol{E}$ in Eq. (9.4) and substituting the Lorentz condition

$$\boldsymbol{\nabla}\cdot\boldsymbol{A} = -\frac{1}{c}\frac{\partial\phi}{\partial t}, \tag{9.14}$$

Eq. (9.4) reduces to

$$\frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} - \boldsymbol{\nabla}^2\phi = 4\pi\rho. \tag{9.15}$$

Equations (9.13) and (9.15) can be combined and written in a relativistically covariant form

$$\left(\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2\right)A^\mu = \frac{4\pi j^\mu}{c}; \qquad (\mu = 0, 1, 2, 3) \tag{9.16}$$

using the four-vector notation. A four-vector is defined as a set of four quantities that transform in the same way as ct, x, y, z when one goes from one inertial frame of reference to another (under Lorentz transformation).

$$\begin{array}{lll} x^\mu & (= x^0, x^1, x^2, x^3); & x^0 = ct; \\ A^\mu & (= A^0, A^1, A^2, A^3); & A^0 = \phi; \\ j^\mu & (= j^0, j^1, j^2, j^3); & j^0 = c\rho. \\ p^\mu & (= p^0, p^1, p^2, p^3); & p^0 = E/c. \end{array} \tag{9.17}$$

Defining the D'Alembertian operator

$$\Box = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2,$$

Eq. (9.16) becomes

$$\Box A^\mu = \frac{4\pi j^\mu}{c}. \tag{9.18}$$

Equation (9.18) can be considered as the relativistic wave equation for the photon but the A^μ's are subject to the Lorentz restriction (9.14).

$$\sum_\mu \frac{\partial A^\mu}{\partial x^\mu} = 0 \qquad \text{or} \qquad \sum_\mu \partial_\mu A^\mu = 0, \tag{9.19}$$

using the short-hand notation $\partial_\mu = \partial/\partial x^\mu$. Potentials satisfying the Lorentz condition, Eq. (9.19) are known to be in Lorentz gauge. It is to be observed that Eq. (9.18) together with the Lorentz condition (9.19) is equivalent to Maxwell's equations. Although A^μ is often referred to as potential in classical physics, it really refers to the photon field in field theory.

For a free field, $\boldsymbol{j}^\mu = 0$. Consequently, the electromagnetic wave equation reduces to

$$\Box A^\mu = 0 \qquad \text{or} \qquad \sum_{\nu=0}^{3} \partial^\nu \partial_\nu A^\mu = 0. \tag{9.20}$$

On inspection, it can be recognized as the Klein-Gordon equation for a particle with zero mass. The solution which we sought for the Klein-Gordon equation was a scalar function corresponding to spin zero particle. In the present case, we have to find a solution for the four-vector function and this will correspond to the spin 1 particle.

9.2 Electromagnetic Field Tensor

Given the scalar and vector potentials ϕ and $\boldsymbol{A}$, they can be treated as the components A^μ of a four-component vector potential and the components of the magnetic and electric fields can be expressed as the components of a second rank antisymmetric tensor $F^{\mu\nu}$, defined by[1]

$$F^{\mu\nu} = \frac{\partial A^\mu}{\partial x_\nu} - \frac{\partial A^\nu}{\partial x_\mu} = \partial^\nu A^\mu - \partial^\mu A^\nu. \tag{9.21}$$

using the notation $\partial^\nu = \partial/\partial x_\nu$. Since this second rank tensor is antisymmetric $F^{\mu\nu} = -F^{\nu\mu}$ and the diagonal elements with $\mu = \nu$ are zero, there are only six independent components and it can be shown that they correspond to the three components of E and the three components of B.

From Eqs. (9.5) and (9.8), we obtain

$$B_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} \equiv -\partial^2 A^3 + \partial^3 A^2 = F^{23}; \tag{9.22}$$

$$B_y = \frac{\partial A_x}{\partial z} - \frac{\partial A_z}{\partial x} \equiv -\partial^3 A^1 + \partial^1 A^3 = F^{31}; \tag{9.23}$$

$$B_z = \frac{\partial A_y}{\partial x} - \frac{\partial A_x}{\partial y} \equiv -\partial^1 A^2 + \partial^2 A^1 = F^{12}; \tag{9.24}$$

$$E_x = -\frac{1}{c}\frac{\partial A_x}{\partial t} - \frac{\partial \phi}{\partial x} \equiv -\partial^0 A^1 + \partial^1 A^0 = F^{01}; \tag{9.25}$$

$$E_y = -\frac{1}{c}\frac{\partial A_y}{\partial t} - \frac{\partial \phi}{\partial y} \equiv -\partial^0 A^2 + \partial^2 A^0 = F^{02}; \tag{9.26}$$

$$E_z = -\frac{1}{c}\frac{\partial A_z}{\partial t} - \frac{\partial \phi}{\partial z} \equiv -\partial^0 A^3 + \partial^3 A^0 = F^{03}. \tag{9.27}$$

Collecting the above results, we can explicitly write down the elements of the field tensor in the form of a matrix, designating the rows and columns by μ and ν.

[1] The differential operators are defined by

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = (\partial_0, \partial_1, \partial_2, \partial_3) = \frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$$

$$= \frac{1}{c}\frac{\partial}{\partial t}, \boldsymbol{\nabla} \quad .$$

$$\partial^\mu = g^{\mu\nu}\partial_\nu = \frac{1}{c}\frac{\partial}{\partial t}, -\boldsymbol{\nabla} \quad .$$

$$F^{\mu\nu} = \begin{array}{c} \mu/\nu\rightarrow \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{bmatrix} \end{array}. \tag{9.28}$$

In a crisp way, the field tensor can be written as[2]

$$F^{\mu\nu}\,\delta_{\mu\nu} = 0; \qquad F^{ij} = \epsilon_{ijk}\,B_k; \qquad F^{0k} = -F^{k0} = E_k, \tag{9.29}$$

Maxwell's equations (9.3) and (9.4) can be combined and written in terms of the electromagnetic field tensor $F^{\mu\nu}$.

$$\sum_{\mu=0}^{3} \partial_\mu F^{\mu\nu} = \frac{4\pi j^\nu}{c}, \qquad \nu = 0, 1, 2, 3. \tag{9.30}$$

For a free field, $j^\nu = 0$. Consequently,

$$\sum_{\mu} \partial_\mu F^{\mu\nu} \equiv \sum_{\mu} \frac{\partial F^{\mu\nu}}{\partial x^\mu} = 0, \qquad \text{for a free field.} \tag{9.31}$$

Eqs. (9.30) are equivalent to the inhomogeneous set of Maxwell's equations (9.3) and (9.4). The other set (*viz.*, the homogeneous set) of Maxwell's equations (9.1) and (9.2) follows directly from the definition (9.21) of field tensor $F^{\mu\nu}$.

$$\frac{\partial F^{\mu\nu}}{\partial x_\lambda} + \frac{\partial F^{\nu\lambda}}{\partial x_\mu} + \frac{\partial F^{\lambda\mu}}{\partial x_\nu} = 0, \tag{9.32}$$

where μ, ν, λ are any three cyclic set out of the four indices 0,1,2,3. When μ, ν, λ are chosen as 1,2,3, Eq. (9.32) corresponds to Maxwell's equation $\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0$. Choosing the indices μ, ν, λ as (2,3,0), (3,0,2), (0,2,3) in Eq. (9.32), we obtain the other homogeneous Maxwell's equation (9.2).

Thus, we find that the very definition of the field tensors $F^{\mu\nu}$ satisfy the homogeneous set of Maxwell's equations and so we need to consider

[2]The symbol ϵ_{ijk} stands for

$$\begin{aligned} \epsilon_{ijk} &= +1, && \text{if } i,j,k \text{ is an even permutation of the numbers 1,2,3} \\ &= -1, && \text{if } i,j,k \text{ is an odd permutation of the numbers 1,2,3} \\ &= 0, && \text{if any two indices are equal.} \end{aligned}$$

only the inhomogeneous equations (9.30) for seeking a Lagrangian formulation.

Treating A_μ as field quantities, the Lagrangian density for the electromagnetic field can be written as

$$\mathscr{L} = -\frac{F_{\lambda\rho}F^{\lambda\rho}}{16\pi} + \frac{j_\lambda A^\lambda}{c}. \tag{9.33}$$

The first term on the right hand side of Eq. (9.33) is for the free field and the second term denotes the interaction in terms of the four-vector current density j_μ. It can be shown that the inhomogeneous Maxwell's Eqs. (9.30) can be obtained from the Lagrangian density (9.33) by insisting that they satisfy the Euler-Lagrange Equation,

$$\frac{\partial \mathscr{L}}{\partial A_\mu} - \partial_\nu \frac{\partial \mathscr{L}}{\partial(\partial_\nu A_\mu)} = 0. \tag{9.34}$$

It follows from (9.33) that

$$\frac{\partial \mathscr{L}}{\partial A_\mu} = \frac{j_\lambda}{c}\delta_{\lambda\mu} \tag{9.35}$$

$$\frac{\partial \mathscr{L}}{\partial\left(\frac{\partial A_\mu}{\partial x_\nu}\right)} = -\frac{F_{\lambda\rho}}{8\pi}\frac{\partial F_{\lambda\rho}}{\partial\left(\frac{\partial A_\mu}{\partial x_\nu}\right)} \tag{9.36}$$

Since

$$F_{\lambda\rho} = \frac{\partial A_\rho}{\partial x_\lambda} - \frac{\partial A_\lambda}{\partial x_\rho},$$

we obtain

$$\sum_{\lambda\rho} F_{\lambda\rho}\frac{\partial F_{\lambda\rho}}{\partial\left(\frac{\partial A_\mu}{\partial x_\nu}\right)} = \sum_{\lambda\rho} F_{\lambda\rho}\left(\delta_{\lambda\mu}\delta_{\rho,\nu} - \delta_{\lambda\nu}\delta_{\rho,\mu}\right) = 2F_{\mu\nu}. \tag{9.37}$$

Substituting these results in the Euler-Lagrange Eq. (9.34), we get

$$\frac{j_\mu}{c} - \frac{1}{4\pi}\frac{\partial F_{\mu\nu}}{\partial x_\nu} = 0, \tag{9.38}$$

which is identical with Eq. (9.30).

The Lagrangian density $\mathscr{L}$ as given by Eq. (9.33 is a sum of two terms, one arising from the free field and the other from the interaction with the outside world in the form of charge and current density. The charge and current density is an explicit function of the spatial and time coordinates

x_μ and so only in the case of free fields, there will be conserved currents in the form of stress-energy tensor $T_{\mu\nu}$, defined by

$$\begin{aligned} T_{\mu\nu} &= \frac{\partial \mathscr{L}}{\partial\left(\frac{\partial A_\lambda}{\partial x_\nu}\right)} \frac{\partial A\lambda}{\partial x_\mu} - \mathscr{L}\delta_{\mu\nu} \\ &= -\frac{1}{4\pi} F_{\nu\lambda} A_{\lambda\mu} - \mathscr{L}\delta_{\mu\nu}. \end{aligned} \tag{9.39}$$

It is possible to study the electromagnetic theory without using the potentials at all but using only the field strengths. This will be more satisfactory since the field strengths are the only physically measurable quantities and not the potentials. But the use of potentials greatly simplifies the calculations and becomes a necessity when one wants to treat the interaction and quantization of the electromagnetic field.

9.3 Quantization of the Electromagnetic Field

The quantization of the free field[3] equation (9.20) is beset with certain difficulties.

1. Equation (9.20) is equivalent to Maxwell's equations only if it is considered in conjunction with the Lorentz condition (9.19). On the other hand, the canonical quantization procedure, followed hitherto, is applicable only to the independent fields. In the present case, there are four fields but all of them are not free and they are subject to Lorentz restriction (9.19).

2. To maintain the Lorentz invariance of the four-vector field A_μ, the introduction of the matrix tensor $g_{\mu\nu}$ imposes opposite signs to the zeroth component A_0 and the other three components A_1, A_2, A_3. This makes it difficult to treat all the four A_μ's on the same footing.

First let us treat all the four A_μ's as independent and Hermitian and ignore the difference in sign of A_0 and the Lorentz constraint (9.19). The difficulties 1 and 2 are overcome by a procedure, prescribed by Gupta and Bleuler[4].

[3]The electromagnetic field in a space region which is far removed from any sources of the field is also called *the radiation field*.

[4]S.N. Gupta, *Proc. Phys. Soc.* **A63**, 681 (1950); K. Bleuler, *Helv. Phys. Acta* **23**, 567 (1950).

The plane wave solution of (9.20) is given by[5]

$$u_\lambda(\mathbf{x}) = \frac{1}{\sqrt{V}}\frac{1}{\sqrt{2\omega_k}}\epsilon^{(\lambda)}(\boldsymbol{k})e^{-i\mathbf{k}\cdot\mathbf{x}}, \qquad (\lambda = 0,1,2,3) \tag{9.40}$$

with

$$\mathbf{k}\cdot\mathbf{x} = \omega_k t - \boldsymbol{k}\cdot\boldsymbol{x}; \tag{9.41}$$

$$\mathbf{k}^2 = \omega_k^2 - \boldsymbol{k}^2 = 0, \qquad \omega_k = |\boldsymbol{k}| > 0. \tag{9.42}$$

In Eq. (9.40), $\epsilon^{(\lambda)}(\boldsymbol{k})$ $(\lambda = 0,1,2,3)$ denote a complete set of orthonormal vectors in the Minkowski space for each value of $\boldsymbol{k}$. The field operator $A_\mu(\mathbf{x})$ can be expanded in terms of $u_\lambda(\mathbf{x})$.

$$A_\mu(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\boldsymbol{k}}\frac{1}{\sqrt{2\omega_k}}\sum_{\lambda=0}^{3}\epsilon_\mu^{(\lambda)}(\boldsymbol{k})\left[a^{(\lambda)}(\boldsymbol{k})e^{-i\mathbf{k}\cdot\mathbf{x}} + a^{(\lambda)\dagger}(\boldsymbol{k})e^{i\mathbf{k}\cdot\mathbf{x}}\right] \tag{9.43}$$

In Eq. (9.43), $\epsilon_\mu^{(\lambda)}(\boldsymbol{k})$, $\lambda = 0,1,2,3$ represent unit vectors in the four-dimensional space which are normalized in a Lorentz invariant way, $\epsilon^{(0)}$ being time-like and $\epsilon^{(1)}, \epsilon^{(2)}, \epsilon^{(3)}$, space-like.

$$\begin{aligned}\epsilon^{(\lambda)}\cdot\epsilon^{(\lambda)} &= \epsilon_\mu^{(\lambda)}\epsilon^{(\lambda')}{}_\mu = \epsilon_\mu^{(\lambda)}g^{\mu\nu}\epsilon_\nu^{(\lambda')} \\ &= \epsilon_0^{(\lambda)}\epsilon_0^{(\lambda')} - \epsilon_1^{(\lambda)}\epsilon_1^{(\lambda')} - \epsilon_2^{(\lambda)}\epsilon_2^{(\lambda')} - \epsilon_3^{(\lambda)}\epsilon_3^{(\lambda')} \\ &= g^{\lambda\lambda'}.\end{aligned} \tag{9.44}$$

Let us take advantage of the arbitrariness of the vector $\epsilon^{(\lambda)}$ and choose $\epsilon^{(3)}(\boldsymbol{k})$ along the direction of $\boldsymbol{k}$ and the vectors $\epsilon^{(1)}(\boldsymbol{k})$ and $\epsilon^{(2)}(\boldsymbol{k})$ perpendicular to $\boldsymbol{k}$. Further, the x_3 axis is chosen along $\boldsymbol{k}$. With this choice, $\epsilon^{(1)}$, $\epsilon^{(2)}$ denote the transverse polarization, $\epsilon^{(3)}$, the longitudinal polarization and $\epsilon^{(0)}$, the scalar or time-like polarization. The scalar and longitudinal photons are unphysical and only the transverse photons are allowed in nature. Let us define the following quantities

$$A_\mu(\boldsymbol{k}) = \sum_{\lambda=0}^{3}\epsilon_\mu^{(\lambda)}(\boldsymbol{k})a^{(\lambda)}(\boldsymbol{k}); \qquad A_\mu^\dagger(\boldsymbol{k}) = \sum_{\lambda=0}^{3}\epsilon_\mu^{(\lambda)}(\boldsymbol{k})a^{(\lambda)\dagger}(\boldsymbol{k}). \tag{9.45}$$

In terms of these quantities, the field variable $A_\mu(\mathbf{x})$ can be written as

$$A_\mu(\mathbf{x}) = \frac{1}{\sqrt{V}}\sum_{\boldsymbol{k}}\frac{1}{\sqrt{2\omega_k}}\left[A_\mu(\boldsymbol{k})e^{-i\mathbf{k}\cdot\mathbf{x}} + A_\mu^\dagger(\boldsymbol{k})e^{i\mathbf{k}\cdot\mathbf{x}}\right] \tag{9.46}$$

$$= A_\mu^{(+)}(\mathbf{x}) + A_\mu^{(-)}(\mathbf{x}), \tag{9.47}$$

[5]The four-vectors are denoted by upright bold face letters like $\mathbf{k}$, $\mathbf{x}$ and the three vectors by italic bold face letters like $\boldsymbol{k}$, $\boldsymbol{x}$.

where $A_\mu^{(+)}(\mathbf{x})$ and $A_\mu^{(-)}(\mathbf{x})$ represent the positive and negative frequency parts of the field operator $A_\mu(\mathbf{x})$.

The operators $A_\mu(\boldsymbol{k})$ and $A_\nu^\dagger(\boldsymbol{k}')$ obey the following commutation relations.

$$[A_\mu(\boldsymbol{k}), A_\nu^\dagger(\boldsymbol{k}')]_- = -g_{\mu\nu}\delta(\boldsymbol{k}-\boldsymbol{k}'), \tag{9.48}$$

where $g_{\mu\nu}$ denotes the metric that we use in representing the four-vector.

The corresponding commutation relations for the annihilation and creation operators $a^{(\lambda)}$ and $a^{(\lambda)\dagger}$. are

$$\left[a^{(\lambda)\dagger}(\boldsymbol{k}), a^{(\lambda')\dagger}(\boldsymbol{k}')\right]_- = \left[a^{(\lambda)}(\boldsymbol{k}), a^{(\lambda')}(\boldsymbol{k}')\right]_- = 0; \tag{9.49}$$

$$\left[a^{(\lambda)}(\boldsymbol{k}),\ a^{(\lambda')\dagger}(\boldsymbol{k}')\right]_- = -g_{\lambda\lambda'}\,\delta(\boldsymbol{k}-\boldsymbol{k}'). \tag{9.50}$$

Let us write more explicitly Eq. (9.50).

$$\left[a^{(0)}(\boldsymbol{k}), a^{(0)\dagger}(\boldsymbol{k}')\right]_- = -\delta(\boldsymbol{k}-\boldsymbol{k}'); \tag{9.51}$$

$$\left[a^{(i)}(\boldsymbol{k}),\ a^{(j)\dagger}(\boldsymbol{k}')\right]_- = \delta_{ij}\,\delta(\boldsymbol{k}-\boldsymbol{k}'), \qquad i,j=1,2,3. \tag{9.52}$$

The interpretation of a and $a^\dagger$ as annihilation and creation operators for $\lambda = 1,2,3$, that is for longitudinal and transverse photons, poses no problems but the relation for scalar photons will give problems because of the minus sign on the right hand side of Eq. (9.51). One problem is with respect to the norm of the scalar photon state and the other problem is with respect to the negative energy contribution by the scalar photons. Following the procedure analogous to the one adopted in the case of scalar fields, one can get the following expression for the Hamiltonian.

$$H = \sum_k \omega_k \left[\sum_{\lambda=1}^{3} a^{(\lambda)\dagger}(\boldsymbol{k})a^{(\lambda)}(\boldsymbol{k}) - a^{(0)\dagger}(\boldsymbol{k})a^{(0)}(\boldsymbol{k})\right]. \tag{9.53}$$

Actually, there is a subtle difference. The number density operator for the time-like photon (scalar photon) is not $a^{(0)\dagger}(\boldsymbol{k})a^{(0)}(\boldsymbol{k})$ but $-a^{(0)\dagger}(\boldsymbol{k})a^{(0)}(\boldsymbol{k})$.

9.4 The Gupta-Bleuler Formulation

The above quantization procedure suffers from two defects.

1. It ignores the Lorentz condition.

$$\partial_\mu A_\mu(\mathbf{x}) = 0.$$

2. It assumes that all the field operators $A_\mu(\mathbf{x})$ are Hermitian. If they are so, then the vacuum expectation value of $A_\mu(\boldsymbol{k})A_\nu^\dagger(\boldsymbol{k}')$ is

$$\begin{aligned} \langle \Phi_0, A_\mu(\boldsymbol{k})A_\nu^\dagger(\boldsymbol{k}')\Phi_0\rangle &= \langle A_\mu^\dagger(\boldsymbol{k})\Phi_0, A_\nu^\dagger(\boldsymbol{k}')\Phi_0\rangle \\ &= -g_{\mu\nu}\delta(\boldsymbol{k}-\boldsymbol{k}'), \end{aligned} \tag{9.54}$$

which is positive for $A_j(\boldsymbol{k})A_j^\dagger(\boldsymbol{k}), \quad j = 1,2,3$ and negative for $A_0(\boldsymbol{k})A_0^\dagger(\boldsymbol{k})$. This is an inconsistency within the theory. It is not possible to exclude the time-like photons (scalar photons) since they are required for the formulation of a relativistic invariant theory.

The above defects were rectified by Gupta and Bleuler (1) by insisting that only the expectation values of the operators need to comply with the classical relations but not the operators themselves and (2) by redefining the scalar product using a metric operator η.

The norm $|\Phi|$ of a state Φ is usually defined by

$$|\Phi| = \langle \Phi|\Phi\rangle. \tag{9.55}$$

This is redefined using an indefinite metric operator η.

$$|\Phi| = \langle \Phi|\eta|\Phi\rangle. \tag{9.56}$$

Since the norm should be real, it follows that

$$\eta = \eta^\dagger.$$

The definition of the norm of a vector is somewhat relaxed. It needs no longer be positive definite. It can be positive, negative or zero. But only the first of these possibilities is physically meaningful and admits the probability interpretation.

With this modified definition of norm, the expectation value of an operator A becomes

$$\langle A\rangle = \langle \Phi|\eta A|\Phi\rangle. \tag{9.57}$$

Le us choose η in such a way that it commutes with $A_j(\mathbf{x})$, $j = 1,2,3$ but anticommutes with $A_0(\mathbf{x})$.

$$\begin{aligned} [\eta, A_j(\mathbf{x})]_- &= 0, \qquad j = 1,2,3. \\ \{\eta, A_0(\mathbf{x})\}_+ &= 0. \end{aligned} \tag{9.58}$$

Also

$$\begin{aligned} \left[\eta^2, A_j\right]_- &= \eta\left[\eta, A_j\right]_- + \left[\eta, A_j\right]_- \eta = 0, \qquad j = 1,2,3. \\ \left[\eta^2, A_0\right]_- &= \eta\left\{\eta, A_0\right\}_+ - \left\{\eta, A_0\right\}_+ \eta = 0. \end{aligned} \tag{9.59}$$

Thus η^2 is a c-number that commutes with all the operators of the field and so can be chosen to yield unity.

$$\eta^2 = \eta\eta^\dagger = 1.$$

9.4.1 The indefinite metric operator η

The metric operator η is so chosen to obey the commutation relations (9.58) with the field operator A_μ. From Eq. (9.46), one can obtain the commutation relations with the annihilation and creation operators[6].

$$\begin{aligned} [\eta, a_j(\boldsymbol{k})] &= 0, \qquad j = 1,2,3. \\ \{\eta, a_0(\boldsymbol{k})\} &= 0. \end{aligned} \tag{9.60}$$

Let us consider, for simplicity, a state Φ which consists of photons of momentum $\boldsymbol{k}$ only. Then, in the occupation number representation, the state Φ can be represented as

$$\Phi = |n_0, n_1, n_2, n_3\rangle, \qquad n_\lambda = n_\lambda(\boldsymbol{k}).$$

It is a simple exercise to obtain the matrix elements of $a_\lambda(\boldsymbol{k})$ and $a_\lambda^\dagger(\boldsymbol{k})$.

$$\langle n_0', n_1', n_2', n_3'|a_\lambda(\boldsymbol{k})|n_0, n_1, n_2, n_3\rangle = \sqrt{n_\lambda(\boldsymbol{k})}\,\delta_{n_\lambda', n_\lambda - 1} \prod_{i\neq\lambda} \delta_{n_i, n_i'}. \tag{9.61}$$

$$\langle n_0, n_1, n_2, n_3|a_\lambda^\dagger(\boldsymbol{k})|n_0', n_1', n_2', n_3'\rangle = \sqrt{n_\lambda(\boldsymbol{k})}\,\delta_{n_\lambda, n_\lambda' + 1} \prod_{i\neq\lambda} \delta_{n_i, n_i'}. \tag{9.62}$$

The commutation relations (9.60) can be explicitly written as

$$\begin{aligned} &\langle n_0', n_1', n_2', n_3'|\eta a_j(\boldsymbol{k})|n_0, n_1, n_2, n_3\rangle \\ &\quad = \langle n_0', n_1', n_2', n_3'|a_j(\boldsymbol{k})\eta|n_0, n_1, n_2, n_3\rangle, \qquad j = 1,2,3. \end{aligned} \tag{9.63}$$

$$\begin{aligned} &\langle n_0', n_1', n_2', n_3'|\eta a_4(\boldsymbol{k})|n_0, n_1, n_2, n_3\rangle \\ &\quad = -\langle n_0', n_1', n_2', n_3'|a_0(\boldsymbol{k})\eta|n_0, n_1, n_2, n_3\rangle. \end{aligned} \tag{9.64}$$

[6]The annihilation and creation operators are represented either as $a^{(\lambda)}$, $a^{(\lambda)\dagger}$ or as a_λ, $a_\lambda^\dagger$, using λ as a superscript or as a subscript, depending upon convenience.

From Eqs. (9.61)–(9.64), we infer that η is a diagonal matrix whose matrix elements are explicitly given below:

$$\begin{aligned} \langle \cdots n_j \cdots |\eta| \cdots n_j \cdots \rangle &= \langle \cdots n_j + 1 \cdots |\eta| \cdots n_j + 1 \cdots \rangle, \\ & \qquad j = 1, 2, 3. \qquad (9.65) \\ \langle n_0 \cdot \cdots \cdot |\eta| n_0 \cdot \cdots \cdot \rangle &= -\langle n_0 + 1 \cdot \cdots \cdot |\eta| n_0 + 1 \cdot \cdots \cdot \rangle. \qquad (9.66) \end{aligned}$$

It follows that the square of the matrix element is unity.

$$|\langle n_0, n_1, n_2, n_3 | \eta | n_0, n_1, n_2, n_3 \rangle|^2 = 1 \qquad (9.67)$$

From the above discussion, we conclude that η is a diagonal matrix with the matrix element

$$\langle n_0', n_1', n_2', n_3' | \eta | n_0, n_1, n_2, n_3 \rangle = (-1)^{n_0} \prod_{\lambda=0}^{3} \delta_{n_\lambda, n_\lambda'}. \qquad (9.68)$$

If the state Φ consists of photons of several momenta $\boldsymbol{k}$, then

$$\langle \Phi | \eta | \Phi \rangle = (-1)^{S_k} \prod_{\boldsymbol{k}} \prod_{\lambda=0}^{3} \delta_{n_\lambda(\boldsymbol{k}), n_\lambda'(\boldsymbol{k})}, \qquad (9.69)$$

with

$$S_k = \sum_{\boldsymbol{k}} n_0(\boldsymbol{k}).$$

9.4.2 The Lorentz condition

The Lorentz condition explains the observed transversality of the electromagnetic waves. For the quantum theory to have the correct classical limits, it is sufficient if the expectation values of the operators and not the operators themselves obey the classical equations.

The Lorentz condition is

$$\partial^\mu A_\mu = 0.$$

It is sufficient if the expectation value obeys the Lorentz condition.

$$\langle \partial^\mu A_\mu \rangle = \langle \Phi | \eta \, \partial^\mu A_\mu | \Phi \rangle = 0. \qquad (9.70)$$

Expanding A_μ in terms of positive and negative frequency parts as shown in Eq. (9.47), we get

$$\langle \Phi | \eta \, \partial^\mu A_\mu | \Phi \rangle = \langle \Phi | \eta \, \partial^\mu A_\mu^{(+)} | \Phi \rangle + \langle \Phi | \eta \, \partial^\mu A_\mu^{(-)} | \Phi \rangle \qquad (9.71)$$

Since

$$\langle \Phi, \eta\, \partial^\mu A_\mu^{(-)} \Phi \rangle = \langle \eta\, \partial^\mu A_\mu^{(+)} \Phi, \Phi \rangle, \tag{9.72}$$

it is enough if the subsidiary condition

$$\sum_\mu \partial^\mu A_\mu^{(+)} \Phi = 0 \tag{9.73}$$

is obeyed. This yields the allowed states that satisfies the Lorentz condition. Let us proceed to find out the allowed states. From Eqs. (9.46) and (9.47), we get

$$A_\mu^{(+)}(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{k}} \frac{1}{\sqrt{2\omega_k}} a_\lambda(\boldsymbol{k}) \epsilon_\mu^{(\lambda)}(\boldsymbol{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}; \tag{9.74}$$

$$\partial_\mu A_\mu^{(+)}(\mathbf{x}) = \frac{1}{\sqrt{V}} \sum_{\boldsymbol{k}} \frac{-ik_\mu}{\sqrt{2\omega_k}} a_\lambda(\boldsymbol{k}) \epsilon_\mu^{(\lambda)}(\boldsymbol{k}) e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{9.75}$$

We have chosen the photon momentum $\boldsymbol{k}$ and the polarization direction $\epsilon^{(3)}$ along the x_3 axis. So, $\epsilon^{(1)}$ and $\epsilon^{(2)}$ denote the transverse polarizations. With this choice,

$$k_1 = k_2 = 0; \quad k_3 = -|\boldsymbol{k}| = -\omega_k; \quad k_0 = \omega_k.$$

In the coordinate system, so chosen, the subsidiary condition yields

$$\sum_{\mu=0}^{3} \partial^\mu A_\mu^{(+)}(\mathbf{x}) = \frac{-i}{\sqrt{2V}} \sum_{\boldsymbol{k}} \sqrt{\omega_k} \left(a^{(0)}(\boldsymbol{k}) - a^{(3)}(\boldsymbol{k}) \right) e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{9.76}$$

The restriction is on the allowed states Φ of the field rather than a condition on the operators A_μ, since Eq. (9.73) leads to the condition

$$\left(a^{(3)}(\boldsymbol{k}) - a^{(0)}(\boldsymbol{k}) \right) \Phi = 0. \tag{9.77}$$

The restriction is on the combination of longitudinal and scalar photons and does not affect the transverse photons. It is possible to construct the same set of transverse photons but different combinations of longitudinal and scalar photons in accordance with Eq. (9.76). The different states so constructed are related by gauge transformations. For a free field, the gauge can be so chosen that only the transverse photons are present.

For free fields (*i.e.* with no charges present), the vacuum can be represented by the state Φ_0, in which no photons of any kind are present. The

vacuum can equally be described by any state that contains no transverse photons but only an equal amount of longitudinal and scalar photons. The latter description will merely correspond to a different choice of Lorentz gauge.

The condition (9.77) makes the Hamiltonian (9.53) of the free fields positive definite, since the free fields in Lorentz gauge contain only transverse photons.

$$H = \sum_k \omega_k (a_1^\dagger a_1 + a_2^\dagger a_2).$$

Review Questions

9.1 Show how Maxwell's equations for electromagnetic waves can be written in terms of scalar and vector potentials ϕ and $\boldsymbol{A}$. Express the electromagnetic wave equation in a relativistic covariant form by defining a four vector potential A^μ. Show that the components of A^μ are not independent and they are constrained by Lorentz condition.

9.2 What are the electromagnetic field tensors? Show how Maxwell's equations can be written in terms of these field tensors.

9.3 What are the difficulties encountered in quantizing the electromagnetic field $\boldsymbol{A}^\mu$ and how are they overcome by the Gupta-Bleuler prescriptions?

Problems

9.1 Given the electromagnetic field tensors $F^{\mu\nu}$ (contravariant tensor of second rank), given by Eq. (9.28), write down explicitly the matrices for $F^{\nu\mu}$ and $F_{\mu\nu}$.

9.2 The electromagnetic field is described by specifying the electrical and magnetic field strengths $\boldsymbol{E}$ and $\boldsymbol{B}$ at each point and their variation with time. It is found convenient to express them in terms of scalar and vector potentials ϕ and $\boldsymbol{A}$ by the following expressions:

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} \qquad \boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}.$$

Show that ϕ and $\boldsymbol{A}$ do not form an unique set but there are more than one set that yield the same value for the physically observable quantities $\boldsymbol{B}$ and $\boldsymbol{E}$.

9.3 Given the Lagrangian density $\mathscr{L}$ of the electromagnetic field A^μ,

$$\mathscr{L} = -\sum_\nu \frac{1}{2}\partial_\nu A^\mu \partial^\nu A_\mu = -\frac{1}{2}\sum_\nu \left(\frac{\partial A^\mu}{\partial x_\nu}\right)^2,$$

obtain the momentum canonically conjugate to A^μ and the Hamiltonian density. Show that the field A^μ satisfies the field equation

$$\Box A^\mu = 0.$$

Solutions to Problems

9.1 Since $F^{\mu\nu}$ is an anti-symmetric matrix,

$$F^{\nu\mu} = -F^{\mu\nu}.$$

$F^{\mu\nu}$ is a contravariant tensor of second rank, whereas $F_{\mu\nu}$ is a covariant tensor of second rank, obtained by lowering the indices.

$$F_{\mu\nu} = g_{\mu\lambda} g_{\nu\rho} F^{\lambda\rho}.$$

$$\begin{array}{lcll} F_{0k} = g_{0\mu} g_{k\nu} F^{\mu\nu} = -F^{0k}, & \mu,\nu = 0,1,2,3; & k = 1,2,3. \\ F_{k0} = g_{k\mu} g_{0\nu} F^{\mu\nu} = -F^{k0}, & \mu,\nu = 0,1,2,3; & k = 1,2,3. \\ F_{kl} = g_{km} g_{ln} F^{mn} = F^{kl}, & k,l,m,n = 1,2,3. \end{array}$$

To obtain the above relations, we have used the following values of the metric tensors:

$$g_{\mu\nu} = 0, \ \ \mu \neq \nu; \qquad g_{00} = 1; \qquad g_{11} = g_{22} = g_{33} = -1.$$

Using the above results, we write down explicitly the related antisymmetric tensors of second rank, given $F^{\mu\nu}$ by Eq. (9.28):

$$F^{\mu\nu} = \begin{array}{c} \mu/\nu\rightarrow \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & B_z & -B_y \\ -E_y & -B_z & 0 & B_x \\ -E_z & B_y & -B_x & 0 \end{bmatrix} \end{array}.$$

$$F^{\nu\mu} = \begin{array}{c} \nu/\mu\rightarrow \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & +B_y \\ E_y & +B_z & 0 & -B_x \\ E_z & -B_y & +B_x & 0 \end{bmatrix} \end{array}.$$

$$F_{\mu\nu} = \begin{array}{c} \mu/\nu\rightarrow \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & B_z & -B_y \\ E_y & -B_z & 0 & B_x \\ E_z & B_y & -B_x & 0 \end{bmatrix} \end{array}.$$

9.2 The scalar and vector potentials ϕ and $\boldsymbol{A}$ are not observable quantities but only the magnetic and electrical field strengths are observable and physically measurable quantities.

$$\boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A}; \qquad \boldsymbol{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}.$$

Since $\mathrm{curl}\,\mathrm{div}\, f(\boldsymbol{x},t) = 0$, where $f(\boldsymbol{x},t)$ is a scalar function, it is possible for two vector potentials $\boldsymbol{A}$ and $\boldsymbol{A}' = \boldsymbol{A} - \mathrm{div} f$ to satisfy the equation for $\boldsymbol{B}$, yielding the same value for $\boldsymbol{B}$.

$$\boldsymbol{B} = \mathrm{curl}\boldsymbol{A} = \mathrm{curl}\boldsymbol{A}',$$

since $\mathrm{curl}\boldsymbol{A}' = \mathrm{curl}\boldsymbol{A} - \mathrm{curl}\,\mathrm{div} f = \mathrm{curl}\boldsymbol{A}$.

For the simultaneous transformation

$$\begin{aligned} \boldsymbol{A} &\longrightarrow \boldsymbol{A}' = \boldsymbol{A} - \boldsymbol{\nabla} f; \\ \phi &\longrightarrow \phi' = \phi + \frac{1}{c}\frac{\partial f}{\partial t}; \end{aligned}$$

we get

$$\begin{aligned} \boldsymbol{B} &= \boldsymbol{\nabla} \times \boldsymbol{A}' = \boldsymbol{\nabla} \times \boldsymbol{A}. \\ \boldsymbol{E} &= -\boldsymbol{\nabla}\phi' - \frac{1}{c}\frac{\partial \boldsymbol{A}'}{\partial t} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\boldsymbol{\nabla}\frac{\partial f}{\partial t} - \frac{1}{c}\frac{\partial}{\partial t}(\boldsymbol{A} - \boldsymbol{\nabla} f) \\ &= -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial \boldsymbol{A}}{\partial t}, \end{aligned}$$

since $\boldsymbol{\nabla}\frac{\partial f}{\partial t} = \frac{\partial}{\partial t}\boldsymbol{\nabla} f$ because the differential operators $\boldsymbol{\nabla}$ and $\frac{\partial}{\partial t}$ commute. In the four-vector notation,

$$A^\mu \ \ (\mu = 0,1,2,3) \ : \ \phi, \boldsymbol{A},$$

and

$$A'^\mu = A^\mu + \frac{\partial f}{\partial x^\mu}.$$

The last equation represents

$$\begin{aligned} \phi' &= \phi + \frac{1}{c}\frac{\partial f}{\partial t} \\ \boldsymbol{A}' &= \boldsymbol{A} - \boldsymbol{\nabla} f. \end{aligned}$$

This property is known as the invariance of electromagnetic field under gauge transformation.

9.3 Given the Lagrangian density $\mathscr{L}$ of the electromagnetic field A^μ,

$$\mathscr{L} = -\sum_\nu \frac{1}{2}\partial_\nu A^\mu \partial^\nu A_\mu = -\frac{1}{2}\sum_\nu \left(\frac{\partial A^\mu}{\partial x^\nu}\right)^2,$$

the momentum canonically conjugate to A^μ is given by

$$\Pi^\mu = \frac{\partial \mathscr{L}}{\partial A^\mu}\partial t = -\frac{\partial A^\mu}{\partial x^0} = -\partial_0 A^\mu.$$

The Hamiltonian density is given by

$$\begin{aligned}
\mathscr{H} &= \Pi_\mu \dot{A}^\mu - \mathscr{L} \\
&= -\left(\frac{\partial A^\mu}{\partial x^0}\right)^2 - \frac{1}{2}\sum_\nu \left(\frac{\partial A^\mu}{\partial x^\nu}\right)^2 \\
&= -\left(\frac{\partial A^\mu}{\partial x^0}\right)^2 + \frac{1}{2}\left(\frac{\partial A^\mu}{\partial x^0}\right)^2 - \frac{1}{2}\sum_k \left(\frac{\partial A^\mu}{\partial x^k}\right)^2 \\
&= -\frac{1}{2}\left(\frac{\partial A^\mu}{\partial x^0}\right)^2 - \frac{1}{2}\sum_k \left(\frac{\partial A^\mu}{\partial x^k}\right)^2
\end{aligned}$$

To deduce the field equation, we need to construct the Euler-Lagrange equation with the given Lagrangian density.

$$\frac{\partial \mathscr{L}}{\partial A^\mu} - \frac{\partial}{\partial x^\nu}\frac{\partial \mathscr{L}}{\partial(\partial_\nu A^\mu)} = 0. \tag{9.78}$$

With the given Lagrangian density, we have

$$\frac{\partial \mathscr{L}}{\partial A^\mu} = 0. \tag{9.79}$$

$$\frac{\partial}{\partial x^\nu}\frac{\partial \mathscr{L}}{\partial(\partial_\nu A^\mu)} = -\frac{\partial}{\partial x^\nu}\frac{\partial A^\mu}{\partial x^\nu} = \frac{\partial^2 A^\mu}{\partial x^{\nu 2}}. \tag{9.80}$$

Substituting (9.79) and (9.80) into Eq. (9.78, we get

$$\frac{\partial^2 A^\mu}{\partial x^{\nu 2}} = 0. \tag{9.81}$$

Equation (9.81) can be written more explicitly as

$$\frac{\partial^2 A^\mu}{\partial x^{0\,2}} - \sum_k \frac{\partial^2 A^\mu}{\partial x^{k 2}} = 0 \qquad \text{or} \qquad \Box A^\mu = 0.$$

It may be mentioned here that all the four fields A^μ, $\mu = 0, 1, 2, 3$ are treated as free fields and the Lorentz condition has not been imposed.

Chapter 10

Interaction between Fields

Till now, we have considered only free fields but all the phenomena that we observe arise from interaction of fields. Interaction between electron and photon fields give rise to a variety of phenomena such as compton effect, electron-positron pair creation, electron-positron pair annihilation, electron-electron scattering, Bremmstrahlung and a host of other processes, that we have discussed in Chap. 4.

It is an extremely difficult problem to solve the coupled non-linear field equations or the field equations with a source term. So, attempts have been made to solve the problem of interacting fields using perturbation theory. The Hamiltonian of the system is divided into that of the free fields plus an interaction Hamiltonian which is treated as a perturbation. This is justifiable if the interaction is sufficiently weak as in the case of electromagnetic interaction with a coupling constant $\alpha \approx 1/137$. This approach has been highly successful in treating the quantum electrodynamic processes.

The Lagrangian density for free fields consists of terms which are quadratic in field functions which are Fourier analyzed in terms of annihilation and creation operators but no perceptible changes will be observed. In the case of interaction between fields, one type of particles will be annihilated and either the same type of particles with different momenta and spin states or another type of particles will be created leading to observable changes in the system. The total lagrangian L and the Lagrangian density $\mathscr{L}$ can be written as a sum of free field Lagrangian $(L_0, \mathscr{L}_0)$ and

interaction Lagrangian $(L_i, \mathscr{L}_i)$.

$$L = L_0 + L_i; \qquad \mathscr{L} = \mathscr{L}_0 + \mathscr{L}_i. \tag{10.1}$$

In a similar way, one can write the Hamiltonian H and the Hamiltonian density $\mathscr{H}$ as a sum of the free-field Hamiltonian and the interacting field Hamiltonian.

$$H = H_0 + H_i; \qquad \mathscr{H} = \mathscr{H}_0 + \mathscr{H}_i. \tag{10.2}$$

Then, it is possible to use the time-dependent perturbation formalism that we have studied in Quantum Mechanics[1] for the development of perturbation theory. The free-field Hamiltonian H_0 can be considered as the unperturbed Hamiltonian and the interaction Hamiltonian H_i can be treated as the perturbation. The interaction picture is found to be most convenient for the development of the perturbation theory.

10.1 The Tomonaga-Schwinger Equation

Consider the state vector Φ that obeys the Schrödinger equation.

$$(H_0 + H_i)\Phi(t) = i\hbar\frac{\partial\Phi(t)}{\partial t}, \tag{10.3}$$

where H_0 denotes the Hamiltonian of the non-interacting fields (*i.e.* free fields) and H_i denotes the interaction between fields. In the Schrödinger picture, the state vector $\Phi(t)$ is time-dependent and the operators H_0, H_i are time-independent. The state vector Φ, in the absence of interaction $(H_i = 0)$ describes a situation in which a set of definite number of free particles with definite momentum and spin travel through space without any interaction with one another. The interaction Hamiltonian H_i causes interaction of these particles with each other and among themselves.

Let

$$\Psi(t) = e^{iH_0t/\hbar}\Phi(t). \tag{10.4}$$

Then differentiating (10.4), we get

$$i\hbar\frac{\partial\Psi(t)}{\partial t} = H_I(t)\Psi(t), \tag{10.5}$$

[1]The reader may refer to chap. 6 and 7 of my book: V. Devanathan, *Quantum Mechanics*, Narosa Publishing House, New Delhi (2005), for a better understanding.

where

$$H_I(t) = e^{iH_0 t/\hbar} H_i e^{-iH_0 t/\hbar} \tag{10.6}$$

is the interaction energy operator in the interaction picture. $H_I(t)$ is explicitly time-dependent in contrast to the Schrödinger operator H_i which is time-independent. In general, the operator $\mathcal{O}_I$ in the interaction picture is related to the Schrödinger operator $\mathcal{O}_S$ by the relation

$$\mathcal{O}_I = e^{iH_0 t/\hbar} \mathcal{O}_S e^{-iH_0 t/\hbar}, \tag{10.7}$$

which, on differentiation, yields the relation

$$i\hbar \frac{\partial \mathcal{O}_I(t)}{\partial t} = [\mathcal{O}_I, H_0]. \tag{10.8}$$

If we take the operator $\mathcal{O} = H_0$, then H_0 is the same in both the pictures, Schrödinger and interaction.

$$(H_0)_{\text{int}} = (H_0)_S.$$

The Hamiltonian H_0 contains only free fields. This assures that the field operators satisfy the free-field equations in the interaction picture also whereas the time-dependence of the state vector $\Psi(t)$ is determined by the interaction energy $H_I(t)$ only.

The interaction picture has distinct advantages. 1. Since the field operators satisfy free field equations, invariant commutation relations can be written down for all times. 2. Equation (10.5) can be generalized to hold in four-dimensional space-time so as to make it co-variant under Lorentz transformation to yield the Tomonaga-Schwinger equation[2].

$$i\hbar c \frac{\delta \Psi(\sigma)}{\delta \sigma(\mathbf{x})} = \mathscr{H}_I(\mathbf{x}) \Psi(\sigma), \tag{10.9}$$

where $\mathscr{H}_I(\mathbf{x})$ denotes the Hamiltonian density at space-time point $\mathbf{x}$ and σ represents the space-like surfaces in the four-dimensional space-time.

10.2 The Invariant Perturbation Theory

The differential equation (10.5) can be written as an integral equation.

$$\Psi(t) = \Psi(t_0) - \frac{i}{\hbar} \int_{t_0}^{t} dt' H_I(t') \Psi(t'). \tag{10.10}$$

[2]S.S. Schweber, H. A. Bethe and F. D. Hoffmann, *Mesons and Fields, Vol 1*, p.167-169, Row, Peterson and Co. (1956).

The integral equation incorporates the initial condition that at $t = t_0$, the state vector $\Psi(t) = \Psi(t_0)$. One can define a time-development operator $U(t, t_0)$ such that

$$\Psi(t) = U(t, t_0)\Psi(t_0), \qquad \Psi(t_0) = U^{-1}(t, t_0)\Psi(t), \tag{10.11}$$

with

$$U(t_0, t_0) = 1.$$

The time-development operator $U(t, t_0)$ is a unitary operator that preserves the normalization of the state vectors $\Psi(t)$. The operator $U(t, t_0)$ satisfies the following differential equation

$$i\hbar \frac{\partial U(t, t_0)}{\partial t} = H_I(t) U(t, t_0), \tag{10.12}$$

the hermitian conjugate of which can be written as

$$-i\hbar \frac{\partial U^\dagger(t, t_0)}{\partial t} = U^\dagger(t, t_0) H_I(t), \tag{10.13}$$

since $H_I(t)$ is Hermitian. It can be easily verified that $U(t, t_0)$ is a unitary operator

$$U(t, t_0) = U^\dagger(t_0, t) = U^{-1}(t_0, t), \tag{10.14}$$

and obeys the group property

$$U(t, t_0) = U(t, t_1) U(t_1, t_0). \tag{10.15}$$

The differential equation (10.12) can be written as an integral equation.

$$\begin{aligned}
\int_{t_0}^{t} \frac{dU(t_1, t_0)}{dt_1} dt_1 &= \frac{1}{i\hbar} \int_{t_0}^{t} H_I(t_1)\, U(t_1, t_0)\, dt_1, \\
U(t, t_0) - U(t_0, t_0) &= \frac{1}{i\hbar} \int_{t_0}^{t} H_I(t_1)\, U(t_1, t_0)\, dt_1, \\
U(t, t_0) &= 1 + \frac{1}{i\hbar} \int_{t_0}^{t} H_I(t_1)\, U(t_1, t_0)\, dt_1.
\end{aligned} \tag{10.16}$$

The last step is obtained by substituting the value $U(t_0, t_0) = 1$. Equation (10.16) is the integral equation for $U(t, t_0)$ and by repeated iteration of $U(t, t_0)$ that occurs within the integral, we obtain an infinite series. Since

$H_I(t)$ is a small perturbation, the successive terms contribute less and less and so the series can be terminated with second or third order justifiably.

$$\begin{aligned} U(t,t_0) &= 1+\frac{1}{i\hbar}\int_{t_0}^{t} H_I(t_1)\,dt_1 \\ &\quad +\left(\frac{1}{i\hbar}\right)^2\int_{t_0}^{t}dt_1\int_{t_0}^{t_1} dt_2 H_I(t_1)\,H_I(t_2)\,U(t_2,t_0) \\ &= 1+\frac{1}{i\hbar}\int_{t_0}^{t} H_I(t_1)\,dt_1 \\ &\quad +\left(\frac{1}{i\hbar}\right)^2\int_{t_0}^{t}dt_1\int_{t_0}^{t_1} dt_2 H_I(t_1)\,H_I(t_2)+\cdots . \end{aligned} \tag{10.17}$$

The infinite series (10.17) can be written in a compact form as shown below.

$$U(t,t_0)=\sum_{n=0}^{\infty} U^{(n)}(t,t_0), \tag{10.18}$$

where

$$U^{(0)}(t,t_0)=1, \tag{10.19}$$

and

$$U^{(n)}(t,t_0)=(i\hbar)^{-n}\int_{t_0}^{t}dt_1\int_{t_0}^{t_1}dt_2\cdots\int_{t_0}^{t_{n-1}}dt_n H_I(t_1)\,H_I(t_2)\cdots H_I(t_n), \tag{10.20}$$

with

$$t_1>t_2>\cdots>t_{n-1}>t_n. \tag{10.21}$$

Dyson[3] observed that this is essentially an integral over the whole time interval from t_0 to t with the restriction that t_j be earlier than t_{j-1}, $j\le n$. One can remove this restriction by introducing the chronological operator P which rearranges a product of time-labeled operators in a chronological time sequence with the latest time occurring first.

$$\begin{aligned} P(H_I(t_1)\,H_I(t_2)\cdots H_I(t_n)) &= H_I(t_i)\,H_I(t_j)\cdots H_I(t_k), \\ & \qquad t_i>t_j>\cdots>t_k. \end{aligned} \tag{10.22}$$

[3]F. J. Dyson, *Phys. Rev.*, **75**, 486, 1736 (1949).

The symbol P denotes *Dyson's Chronological Operator.* Using the symmetry of the integrand and the fact that there are $n!$ possible permutations of the ordering of time, we obtain

$$U^{(n)}(t,t_0) = \frac{(i\hbar)^{-n}}{n!}\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2\cdots\int_{t_0}^{t}dt_n P\left(H_I(t_1)\,H_I(t_2)\cdots H_I(t_n)\right). \quad (10.23)$$

Thus

$$U(t,t_0) = \sum_{n=0}^{\infty}\frac{(i\hbar)^{-n}}{n!}\int_{t_0}^{t}dt_1\int_{t_0}^{t}dt_2\cdots\int_{t_0}^{t}dt_n P\left(H_I(t_1)\,H_I(t_2)\cdots H_I(t_n)\right). \quad (10.24)$$

Equation (10.24) can be rewritten in a form which indicates the covariance better by expressing the interaction Hamiltonian $H_I(t)$ in terms of interaction Hamiltonian density $\mathscr{H}_I(\mathbf{x})$

$$H_I(t) = \int d^3x\mathscr{H}_I(\mathbf{x}),$$

and converting the integral over time into a four-dimensional integral since $\int dtd^3x = \int d^4x/c$.

$$\begin{aligned} U(t,t_0) &= \sum_{n=0}^{\infty}\frac{(i\hbar c)^{-n}}{n!}\int d^4x_1\int d^4x_2\cdots\int d^4x_n \\ &\quad\times P\left(\mathscr{H}_I(\mathbf{x}_1)\,\mathscr{H}_I(\mathbf{x}_2)\cdots\mathscr{H}_I(\mathbf{x}_n)\right), \end{aligned} \quad (10.25)$$

where $\mathbf{x}_1, \mathbf{x}_2, \cdots$ denote the space-time coordinates. Equation (10.25) forms the basis of the modern covariant perturbation theory.

10.3 The S-matrix

The time development operator $U(t,t_0)$ with $t_0 \to -\infty$ and $t \to +\infty$ is often called the $S-$matrix or collision operator. It is also called the scattering matrix.

$$S = U(+\infty,-\infty) \qquad \text{or} \qquad \Psi(+\infty) = S\Psi(-\infty), \quad (10.26)$$

where $\Psi(-\infty)$ and $\Psi(+\infty)$ define asymptotically the incoming and outgoing states.

We usually specify the initial and final states by the eigenfunctions of H_0 since H_I is zero at $t = \pm\infty$. We invoke the adiabatic hypothesis to

slowly switch on the interaction H_I at $t = -\infty$ to allow for self-interaction of the particles although they are well-separated. This process is known as "dressing up" of the particles with their virtual quanta and develop into real physical particles. Then the particles interact and make a transition to the final state. Then the interaction Hamiltonian is slowly switched off and the final state emerges as a different eigenstate of the unperturbed Hamiltonian H_0.

If Φ_i and Φ_f denote the initial and final state of the system, then

$$\begin{aligned} S_{fi} &= \langle \Phi_f | S | \Phi_i \rangle = \langle \Phi_f | U(+\infty, -\infty) | \Phi_i \rangle \\ &= \langle \Phi_f | U(+\infty, 0) U(0, -\infty) | \Phi_i \rangle \\ &= \langle \Psi_f^- | \Psi_i^+ \rangle, \end{aligned} \tag{10.27}$$

where Ψ_i^+ denotes the incoming state and Ψ_f^-, the outgoing state.

10.4 Reduction of the S-matrix

10.4.1 Decomposition into normal products

Each term in the perturbation expansion (10.25) of the S-matrix can give rise to a variety of processes, both virtual as well as real. Dyson[4] and Wick[5] have shown how to express the chronological product

$$P\left(\mathscr{H}_I(\mathbf{x}_1)\, \mathscr{H}_I(\mathbf{x}_2) \cdots \mathscr{H}_I(\mathbf{x}_n)\right)$$

that occurs in the S-matrix in a form that represents explicitly all the physical processes. This is known as the decomposition of the chronological product into normal products. The normal product is defined as the product of free-particle annihilation and creation operators in which all the annihilation operators appear to the right of creation operators. Then, for the given initial and final states, there exists only one normal product that gives a non-zero matrix element between the specified initial and final states. Then the decomposition of S-matrix into normal products is to list all the matrix elements of S in a representation in which all the free particle occupation numbers are diagonal. Then a Feynman diagram is simply a concise way of representing a normal product.

Let us now consider how to reduce the S-matrix into normal products, following the algebraic method of Wick. Consider a product of annihilation and creation operators. To arrange it as a normal product, all the

[4]F. J. Dyson, *Phys. Rev.*, **82**, 428 (1951); **83**, 608 (1951).
[5]G.C. Wick, *Phys. Rev.*, **80**, 268 (1950).

creation operators should be shifted to the left of annihilation operators. This can be done using the commutation relations for scalar field (Boson) operators and anticommutation relations for Dirac field (Fermion) operators, assuming that the commutators and anticommutators are zero. The effect is to include a change in sign when the anticommuting Dirac field operators are shifted from one position to another. Denoting the Scalar field operators by ϕ and the Dirac field operators by ψ and expressing them as a sum of positive and negative frequency parts $\phi = \phi^+ + \phi^-$ and $\psi = \psi^+ + \psi^-$, we get[6]

$$\begin{aligned} N\left(\phi^+(\mathbf{x})\phi^-(\mathbf{y})\right) &= N\left(\phi^-(\mathbf{y})\phi^+(\mathbf{x})\right) \\ &= \phi^-(\mathbf{y})\phi^+(\mathbf{x}). \qquad (10.28) \\ N\left(\psi_\alpha^+(\mathbf{x})\bar{\psi}_\beta^-(\mathbf{y})\right) &= -N\left(\bar{\psi}_\beta^-(\mathbf{y})\psi_\alpha^+(\mathbf{x})\right) \\ &= -\bar{\psi}_\beta^-(\mathbf{y})\psi_\alpha^+(\mathbf{x}). \qquad (10.29) \\ N\left(\psi_\alpha^+(\mathbf{x})\psi_\beta^+(\mathbf{y})\right) &= -N\left(\psi_\beta^+(\mathbf{y})\psi_\alpha^+(\mathbf{x})\right) \\ &= \psi_\alpha^+(\mathbf{x})\psi_\beta^+(\mathbf{y}) = -\psi_\beta^+(\mathbf{y})\psi_\alpha^+(\mathbf{x}). \qquad (10.30) \end{aligned}$$

The normal product operation obeys the distribution law.

$$N\big(\phi^+(\mathbf{x})\left\{\phi^+(\mathbf{y})+\phi^-(\mathbf{z})\right\}\big) = N\big(\phi^+(\mathbf{x})\phi^+(\mathbf{y})\big)+N\big(\phi^+(\mathbf{x})\phi^-(\mathbf{z})\big)\,. \quad (10.31)$$

Using the commutation and anticommutation rules, the definition of normal product and the distribution law, one can arrange the product of any two operators as normal product.

$$\begin{aligned} \phi(\mathbf{x})\phi(\mathbf{y}) &= \left(\phi^+(\mathbf{x}) + \phi^-(\mathbf{x})\right)\left(\phi^+(\mathbf{y}) + \phi^-(\mathbf{y})\right) \\ &= \phi^+(\mathbf{x})\phi^+(\mathbf{y}) + \phi^+(\mathbf{x})\phi^-(\mathbf{y}) + \phi^-(\mathbf{x})\phi^+(\mathbf{y}) + \phi^-(\mathbf{x})\phi^-(\mathbf{y}) \\ &= \phi^+(\mathbf{x})\phi^+(\mathbf{y}) + \phi^-(\mathbf{y})\phi^+(\mathbf{x}) + [\phi^+(\mathbf{x}), \phi^-(\mathbf{y})]_- \\ &\quad +\phi^-(\mathbf{x})\phi^+(\mathbf{y}) + \phi^-(\mathbf{x})\phi^-(\mathbf{y}) \\ &= N\left(\phi^+(\mathbf{x})\phi^+(\mathbf{y})\right) + N\left(\phi^-(\mathbf{x})\phi^+(\mathbf{y})\right) + N\left(\phi^-(\mathbf{x})\phi^+(\mathbf{y})\right) \\ &\quad +N\left(\phi^-(\mathbf{x})\phi^-(\mathbf{y})\right) + [\phi^+(\mathbf{x}), \phi^-(\mathbf{y})]_- \\ &= N\left(\phi(\mathbf{x})\phi(\mathbf{y})\right) + [\phi^+(\mathbf{x}), \phi^-(\mathbf{y})]_- \\ &= N\left(\phi(\mathbf{x})\phi(\mathbf{y})\right) + i\hbar c\Delta^+(\mathbf{x}-\mathbf{y}). \qquad (10.32) \end{aligned}$$

[6]It may be recalled that the positive frequency operators $\phi^+, \psi^+, \bar{\psi}^+$ consist of annihilation operators whereas the negative frequency operators $\phi^-, \psi^-, \bar{\psi}^-$ carry creation operators.

In deducing Eq. (10.32), we have used the relation

$$\phi^+(\mathbf{x})\phi^-(\mathbf{y}) = \phi^-(\mathbf{y})\phi^+(\mathbf{x}) + [\phi^+(\mathbf{x}), \phi^-(\mathbf{y})]_-$$

in the third step and the commutation relation

$$[\phi^+(\mathbf{x}), \phi^-(\mathbf{y})]_- = i\hbar c\Delta^+(\mathbf{x}-\mathbf{y})$$

in the last step.
Similarly, for fermion operators, one gets

$$\begin{aligned}
\bar{\psi}_\alpha(\mathbf{x})\psi_\beta(\mathbf{y}) &= \left(\bar{\psi}^+_\alpha(\mathbf{x}) + \bar{\psi}^-_\alpha(\mathbf{x})\right)\left(\psi^+_\beta(\mathbf{y}) + \psi^-_\beta(\mathbf{y})\right) \\
&= \bar{\psi}^+_\alpha(\mathbf{x})\psi^+_\beta(\mathbf{y}) + \bar{\psi}^+_\alpha(\mathbf{x})\psi^-_\beta(\mathbf{y}) + \bar{\psi}^-_\alpha(\mathbf{x})\psi^+_\beta(\mathbf{y}) \\
&\quad + \bar{\psi}^-_\alpha(\mathbf{x})\psi^-_\beta(\mathbf{y}) \\
&= \bar{\psi}^+_\alpha(\mathbf{x})\psi^+_\beta(\mathbf{y}) - \psi^-_\beta(\mathbf{y})\bar{\psi}^+_\alpha(\mathbf{x}) + \{\bar{\psi}^+_\alpha(\mathbf{x}), \psi^+_\beta(\mathbf{y})\}_+ \\
&\quad + \bar{\psi}^-_\alpha(\mathbf{x})\psi^+_\beta(\mathbf{y}) + \bar{\psi}^-_\alpha(\mathbf{x})\psi^-_\beta(\mathbf{y}) \\
&= N\left(\bar{\psi}^+_\alpha(\mathbf{x})\psi^+_\beta(\mathbf{y})\right) + N\left(\bar{\psi}^+_\alpha(\mathbf{x})\psi^-_\beta(\mathbf{y})\right) \\
&\quad + N\left(\bar{\psi}^-_\alpha(\mathbf{x})\psi^+_\beta(\mathbf{y})\right) + N\left(\bar{\psi}^-_\alpha(\mathbf{x})\psi^-_\beta(\mathbf{y})\right) \\
&\quad + \{\bar{\psi}^+_\alpha(\mathbf{x}), \psi^+_\beta(\mathbf{y})\}_+ \\
&= N\left(\bar{\psi}_\alpha(\mathbf{x})\psi_\beta(\mathbf{y})\right) - iS^-_{\beta\alpha}(\mathbf{y}-\mathbf{x}), \qquad (10.33)
\end{aligned}$$

using the anticommutation relation $\{\bar{\psi}^+_\alpha(\mathbf{x}), \psi^+_\beta(\mathbf{y})\}_+ = -iS^-_{\beta\alpha}(\mathbf{y}-\mathbf{x})$ for Fermion operators.

Using similar algebraic manipulations, we get

$$\begin{aligned}
\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y}) &= N\left(\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y})\right) - iS^+_{\alpha\beta}(\mathbf{x}-\mathbf{y}); \qquad (10.34) \\
\psi(\mathbf{x})\psi(\mathbf{y}) &= N\left(\psi(\mathbf{x})\psi(\mathbf{y})\right); \qquad (10.35) \\
\bar{\psi}(\mathbf{x})\bar{\psi}(\mathbf{y}) &= N\left(\bar{\psi}(\mathbf{x})\bar{\psi}(\mathbf{y})\right); \qquad (10.36) \\
\phi(\mathbf{x})\psi(\mathbf{y}) &= N\left(\phi(\mathbf{x})\psi(\mathbf{y})\right). \qquad (10.37)
\end{aligned}$$

10.4.2 Wick's chronological product

Wick's chronological operator T is similar to Dyson's chronological operator P and arranges the time-labeled field operators $A, B, C, D \cdots$ according to the time sequence, the operators having the earlier times on the right and the operators having the later times on the left. The only difference between these two chronological operators is that Wick's chronological operator includes in its definition the sign of the permutation of

Fermion operators in making such an arrangement.

$$T(ABCD\cdots) = \delta_P P(ABCD\cdots). \tag{10.38}$$

The whole product carries the plus or minus sign δ_P according to the even or odd number of permutations that the Fermion operators have to make in going from one side to the other in the process of time-ordering.

We wish to reiterate that the chronological operator T arranges the Boson and Fermion operators according to the time sequence assuming the commutation and anticommutation relations between these operators to be zero.

$$T\left(\phi(\mathbf{x})\phi(\mathbf{y})\right) = \begin{cases} \phi(\mathbf{x})\phi(\mathbf{y}), & \text{if } x_0 > y_0 \\ \phi(\mathbf{y})\phi(\mathbf{x}), & \text{if } y_0 > x_0. \end{cases} \tag{10.39}$$

$$T\left(\psi(\mathbf{x})\psi(\mathbf{y})\right) = \begin{cases} \psi(\mathbf{x})\psi(\mathbf{y}), & \text{if } x_0 > y_0 \\ -\psi(\mathbf{y})\psi(\mathbf{x}), & \text{if } y_0 > x_0. \end{cases} \tag{10.40}$$

$$T\left(\psi(\mathbf{x})\bar{\psi}(\mathbf{y})\right) = \begin{cases} \psi(\mathbf{x})\bar{\psi}(\mathbf{y}), & \text{if } x_0 > y_0 \\ -\bar{\psi}(\mathbf{y})\psi(\mathbf{x}), & \text{if } y_0 > x_0. \end{cases} \tag{10.41}$$

In the above equations, the suffix 0 indicates that x_0, y_0 are the time components of the space-time coordinates $\mathbf{x}, \mathbf{y}$.

10.4.3 Wick's contractions

The basic problem is how to express Wick's chronological product in terms of a sum of normal products. For this, we introduce the notion of contraction between any two factors, known as Wick's contraction. The contraction represents the commutator or anticommutator bracket that arises from switching the factors in going from a chronological product to normal products.

$$T(AB) = A^{\bullet}B^{\bullet} + N(AB), \tag{10.42}$$

where $A^{\bullet}B^{\bullet}$ represents the contracted pair. Some authors use the symbol of connecting bracket underneath the two operators to denote the contracted pair and write Eq. (10.42) as

$$T(AB) = \underbrace{AB} + N(AB),$$

where $\underbrace{AB}$ denotes the contracted pair, but we shall use the symbol $A^{\bullet}B^{\bullet}$ to denote Wick's contraction. Using Eq. (10.42) as the definition contracted pair, let us consider some simple cases and show that the contracted pair represents a commutation or anticommutation bracket which,

in turn, denotes a propagator.
For $x_0 > y_0$,

$$\begin{aligned}\phi^{+\bullet}(\mathbf{x})\phi^{-\bullet}(\mathbf{y}) &= T\left(\phi^+(\mathbf{x})\phi^-(\mathbf{y})\right) - N\left(\phi^+(\mathbf{x})\phi^-(\mathbf{y})\right)\\ &= \phi^+(\mathbf{x})\phi^-(\mathbf{y}) - \phi^-(\mathbf{y})\phi^+(\mathbf{x})\\ &= \left[\phi^+(\mathbf{x}), \phi^-(\mathbf{y})\right]_-\\ &= i\hbar c\Delta^+(\mathbf{x}-\mathbf{y}). \qquad (10.43)\end{aligned}$$

For $y_0 > x_0$,

$$\begin{aligned}\phi^{+\bullet}(\mathbf{x})\phi^{-\bullet}(\mathbf{y}) &= T\left(\phi^+(\mathbf{x})\phi^-(\mathbf{y})\right) - N\left(\phi^+(\mathbf{x})\phi^-(\mathbf{y})\right)\\ &= \phi^-(\mathbf{y})\phi^+(\mathbf{x}) - \phi^-(\mathbf{y})\phi^+(\mathbf{x})\\ &= 0. \qquad (10.44)\end{aligned}$$

For $x_0 > y_0$,

$$\begin{aligned}\phi^{\bullet}(\mathbf{x})\phi^{\bullet}(\mathbf{y}) &= T\left(\phi(\mathbf{x})\phi(\mathbf{y})\right) - N\left(\phi(\mathbf{x})\phi(\mathbf{y})\right)\\ &= \phi(\mathbf{x})\phi(\mathbf{y}) - N\left(\phi(\mathbf{x})\phi(\mathbf{y})\right)\\ &= i\hbar c\Delta^+(\mathbf{x}-\mathbf{y}), \quad \text{using Eqn. (10.32).} \qquad (10.45)\end{aligned}$$

For $y_0 > x_0$,

$$\begin{aligned}\phi^{\bullet}(\mathbf{x})\phi^{\bullet}(\mathbf{y}) &= T\left(\phi(\mathbf{x})\phi(\mathbf{y})\right) - N\left(\phi(\mathbf{x})\phi(\mathbf{y})\right)\\ &= \phi(\mathbf{y})\phi(\mathbf{x}) - N\left(\phi(\mathbf{x})\phi(\mathbf{y})\right)\\ &= i\hbar c\Delta^+(\mathbf{y}-\mathbf{x}), \quad \text{using Eqn. (10.32).} \qquad (10.46)\end{aligned}$$

In a similar way, we can obtain the results for Fermion contracted pairs.
For $x_0 > y_0$,

$$\begin{aligned}\psi^{\bullet}_{\alpha}(\mathbf{x})\bar{\psi}^{\bullet}_{\beta}(\mathbf{y}) &= T\left(\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y})\right) - N\left(\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y})\right)\\ &= \psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y}) - N\left(\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y})\right)\\ &= -iS^+_{\alpha\beta}(\mathbf{x}-\mathbf{y}), \quad \text{using Eqn. (10.34).} \qquad (10.47)\end{aligned}$$

For $y_0 > x_0$,

$$\begin{aligned}\psi^{\bullet}_{\alpha}(\mathbf{x})\bar{\psi}^{\bullet}_{\beta}(\mathbf{y}) &= T\left(\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y})\right) - N\left(\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y})\right)\\ &= -\bar{\psi}_\beta(\mathbf{y}\psi_\alpha(\mathbf{x})) - N\left(\psi_\alpha(\mathbf{x})\bar{\psi}_\beta(\mathbf{y})\right)\\ &= iS^-_{\alpha\beta}(\mathbf{x}-\mathbf{y}), \quad \text{using Eqn. (10.33).} \qquad (10.48)\end{aligned}$$

Equations (10.47) and (10.48) together yield the Feynman propagator $K_F(\mathbf{x}-\mathbf{y})$.

$$\psi^{\bullet}_{\alpha}(\mathbf{x})\bar{\psi}^{\bullet}_{\beta}(\mathbf{y}) = K_{F\,\alpha\beta}(\mathbf{x}-\mathbf{y}). \qquad (10.49)$$

10.4.4 Wick's theorem

Wick has generalized the Wick's contraction formula (10.42) for a product of a large number of field operators $A(\mathbf{x}_a)B(\mathbf{x}_b)C(\mathbf{x}_c)\cdots Z(\mathbf{x}_z)$.

$$\begin{aligned} T(ABCD\cdots WXYZ) &= N(ABCD\cdots WXYZ) \\ &+N\left(A^{\bullet}B^{\bullet}C\cdots Z\right)+N\left(A^{\bullet}BC^{\bullet}\cdots Z\right)+\cdots+N\left(A\cdots XY^{\bullet}Z^{\bullet}\right) \\ &+N\left(A^{\bullet}B^{\bullet}C^{\bullet\bullet}D^{\bullet\bullet}\cdots WXYZ\right)+\cdots+N(ABCD\cdots W^{\bullet}X^{\bullet}Y^{\bullet\bullet}Z^{\bullet\bullet}) \\ &+\cdots+N\left(A^{\bullet}B^{\bullet}\cdots Y^{\bullet\cdots\bullet}Z^{\bullet\cdots\bullet}\right) \end{aligned} \tag{10.50}$$

The right hand side of Eq. (10.50) consists of normal products with no contractions, with all possible contractions of one pair, two pairs, three pairs and so on. Different contractions are distinguished by using single dot, two dots, three dots and so on as superscripts. Equation (10.50) is known as Wick's theorem. It is proved by the method of induction: If it is true for a product of n operators, it can be shown to be true for $n+1$ operators.

The contractions are mere numbers and not operators. In Eq. (10.50), although they are shown within the normal-ordered products, they can be pulled out of the normal ordered products.

The S-matrix expansion consists of time-ordered product of normal-ordered operators.

$$T\left(\mathscr{H}_I(\mathbf{x}_1)\cdots\mathscr{H}_I(\mathbf{x}_n)\right)=T\left(N(AB\cdots)_{\mathbf{X}_1}\cdots N((AB\cdots)_{\mathbf{X}_n}\right). \tag{10.51}$$

Wick's theorem (10.50) can be extended to include such mixed T-products as shown in Eq. (10.51). In order to avoid commutation or anticommutation between the operators at the same space-time points, the creation operators are given a small increment in time ϵ over the annihilation operators so that the time ordering in each term coincides with normal ordering. At the end, ϵ can be put to zero. This yields

$$\begin{aligned} &T\left(N(AB\cdots)_{\mathbf{X}_1}\cdots N((AB\cdots)_{\mathbf{X}_n}\right) \\ &\quad=T\left((AB\cdots)_{\mathbf{X}_1}\cdots((AB\cdots)_{\mathbf{X}_n}\right)_{\text{no e.t.c.}}, \end{aligned} \tag{10.52}$$

where the right hand side is expanded using Wick's theorem (10.50) with no equal-time contractions (no e.t.c.).

10.5 From S-matrix Expansion to Feynman diagrams

The foregoing discussion on the reduction of the S-matrix into normal products using Wick's theorem will become clearer by considering specific examples in QED. Given the initial and final states, how to pick out the relevant terms in the S-matrix expansion of a given order to write down the transition amplitudes and draw the corresponding Feynman diagrams.

10.5.1 Interaction with external electromagnetic field

Given the interaction term

$$\mathscr{H}_I(\mathbf{x}) = j^\mu(\mathbf{x})A^e_\mu(\mathbf{x}) = \bar{\psi}(\mathbf{x})\gamma^\mu\psi(\mathbf{x})A^e_\mu(\mathbf{x}) = \bar{\psi}(\mathbf{x})\not{A}^e(\mathbf{x})\psi(\mathbf{x}),$$

we can write down the S-matrix expansion as

$$\begin{aligned} S &= 1 + \left(\frac{-ie}{\hbar c}\right)\int d^4x_1 N(\bar{\psi}(\mathbf{x})\not{A}^e(\mathbf{x})\psi(\mathbf{x})) \\ &\quad + \left(\frac{-ie}{\hbar c}\right)^2 \frac{1}{2!}\int d^4x_1 d^4x_2 T\left\{N\left(\bar{\psi}(\mathbf{x}_1)\not{A}^e(\mathbf{x}_1)\psi(\mathbf{x}_1)\right)\right. \\ &\qquad \left. \times N\left(\bar{\psi}(\mathbf{x}_2)\not{A}^e(\mathbf{x}_2)\psi(\mathbf{x}_2)\right)\right\} + \cdots . \end{aligned} \tag{10.53}$$

First order term in S-matrix

Let us consider the first order term in normal order form explicitly in terms of the positive and negative frequency (energy) parts $\psi^-(\mathbf{x})$ and $\psi^+(\mathbf{x})$.

$$\begin{aligned} N(\bar{\psi}(\mathbf{x})\not{A}^e(\mathbf{x})\psi(\mathbf{x})) &= N\left\{(\bar{\psi}^+(\mathbf{x}) + \bar{\psi}^-(\mathbf{x}))\not{A}^e(\mathbf{x})(\psi^+(\mathbf{x}) + \psi^-(\mathbf{x})\right\} \\ &= \bar{\psi}^+(\mathbf{x})\not{A}^e(\mathbf{x})\psi^+(\mathbf{x}) - \psi^-(\mathbf{x})\not{A}^e(\mathbf{x})\bar{\psi}^+(\mathbf{x}) \\ &\quad + \bar{\psi}^-(\mathbf{x})\not{A}^e(\mathbf{x})\psi^+(\mathbf{x}) + \bar{\psi}^-(\mathbf{x})\not{A}^e(\mathbf{x})\psi^-(\mathbf{x}). \end{aligned} \tag{10.54}$$

Let us use the Feynman convention of representing electron by an arrow pointing upwards (forward in time) and positron by an arrow pointing downwards (backward in time) as shown in Fig. 10.1. The positive and negative frequency parts of $\psi(\mathbf{x})$ and $\bar{\psi}(\mathbf{x})$ denote the following events:

$$\begin{array}{lcl} \psi^+(\mathbf{x}) & : & \text{annihilation of electron} \\ \psi^-(\mathbf{x}) & : & \text{creation of positron} \\ \bar{\psi}^+(\mathbf{x}) & : & \text{annihilation of positron} \\ \bar{\psi}^-(\mathbf{x}) & : & \text{creation of electron} \end{array}$$

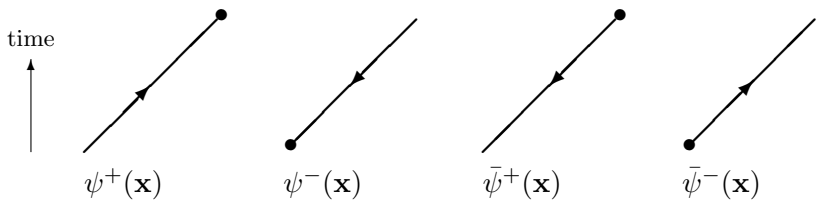


Figure 10.1: Pictorial representation of $\psi^+(\mathbf{x}), \psi^-(\mathbf{x}), \bar{\psi}^+(\mathbf{x}), \bar{\psi}^-(\mathbf{x})$.

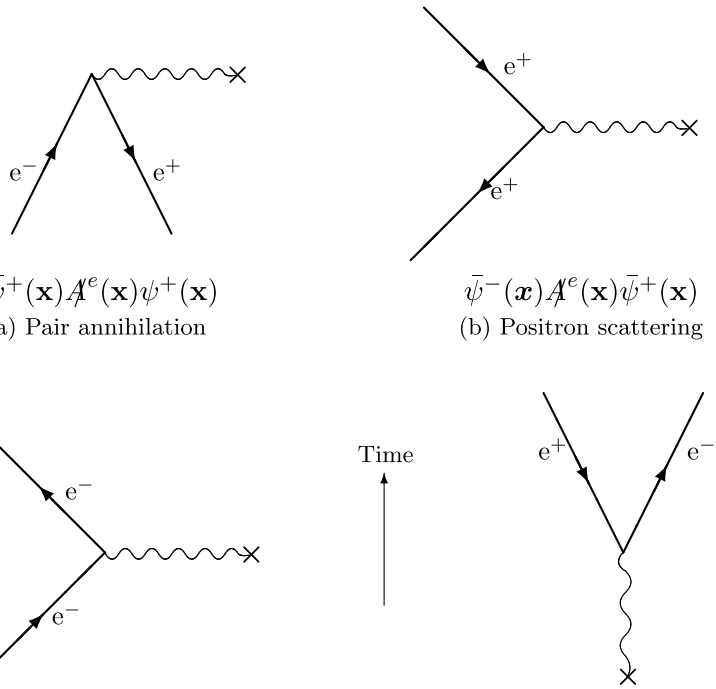
Let us now represent the various terms in Eq. (10.54) by Feynman diagrams as shown in Fig. 10.2. They represent (a) the pair annihilation (b) the positron scattering (c) the electron scattering and (d) the pair creation. These are the basic vertices and all the real processes involving photons, electrons and positrons are built out of them. Energy and momentum are conserved at each vertex but the relativistic energy-momentum relation for all the three particles participating in a vertex will not be simultaneously satisfied. So, all the three particles participating in a vertex will not be real and at least one of them has to be virtual for which the energy-momentum relation need not be satisfied. However, if the external electromagnetic interaction is the Coulomb potential (virtual photon), then the processes (b) and (c) representing positron scattering and electron scattering are possible. For any real process, there must be over-all energy-momentum conservation besides conservation of energy and momentum at each vertex.

We shall now consider the positron scattering by Coulomb potential as an example of the first-order process and show that we get the same scattering amplitude as obtained earlier by Feynman's method. The initial state $|i\rangle$ is specified by a positron of momentum $\boldsymbol{p}_i$ and spin state s_i and the final state $|f\rangle$ by a positron of momentum $\boldsymbol{p}_f$ and spin state s_f.

$$|i\rangle = d^\dagger_{s_i}(\boldsymbol{p}_i)|0\rangle; \qquad |f\rangle = d^\dagger_{s_f}(\boldsymbol{p}_f)|0\rangle.$$

The matrix element for the first order transition from an initial state to a final state is given by

$$\begin{aligned}
\mathcal{M} &= \langle f|S^{(1)}|i\rangle \\
&= -\frac{ie}{\hbar c}\int d^4x_1 \langle f|N(\bar{\psi}(\mathbf{x}_1)\not{A}^e(\mathbf{x}_1)\psi(\mathbf{x}_1))|i\rangle \\
&= -\frac{ie}{\hbar c}\int d^4x_1 \langle 0|d_{s_f}N(\bar{\psi}(\mathbf{x}_1)\not{A}^e(\mathbf{x}_1)\psi(\mathbf{x}_1))d^\dagger_{s_i}|0\rangle. \quad (10.55)
\end{aligned}$$



$\bar{\psi}^+(\mathbf{x})\not{A}^e(\mathbf{x})\psi^+(\mathbf{x})$
(a) Pair annihilation

$\bar{\psi}^-(\boldsymbol{x})\not{A}^e(\mathbf{x})\bar{\psi}^+(\mathbf{x})$
(b) Positron scattering

$\bar{\psi}^-(\boldsymbol{x})\not{A}^e(\mathbf{x})\psi^+(\mathbf{x})$
(c) Electron scattering

$\bar{\psi}^-(\boldsymbol{x})\not{A}^e(\mathbf{x})\psi^-(\mathbf{x})$
(d) Pair creation

Figure 10.2: Feynman diagrams representing basic vertices in QED.

In the expansion of the normal product $N(\bar{\psi}(\mathbf{x}_1)\not{A}^e(\mathbf{x}_1)\psi(\mathbf{x}_1))$, only the term

$$-\psi_\beta^-(\mathbf{x}_1)\not{A}^e\bar{\psi}_\alpha^+(\mathbf{x}_1)$$

will contribute to the process of positron scattering by an external potential $\not{A}^e$. The operator $\bar{\psi}_\alpha^+(\mathbf{x}_1)$ will destroy the positron in the initial state and $\psi_\beta^-(\mathbf{x}_1)$ will create the positron in the final state. Since

$$\psi_\beta^-(\mathbf{x}_1) = \frac{1}{(2\pi)^{3/2}}\int_{p_0} d^3p\frac{1}{\sqrt{2E_p}}\sum_{r=1}^{2} d_r^\dagger(\boldsymbol{p})v_r(\boldsymbol{p})e^{i\mathbf{p}\cdot\mathbf{x}_1}, \tag{10.56}$$

$$\bar{\psi}_\alpha^+(\mathbf{x}_1) = \frac{1}{(2\pi)^{3/2}}\int_{p'_0} d^3p'\frac{1}{\sqrt{2E_{p'}}}\sum_{s=1}^{2} d_s(\boldsymbol{p}')\bar{v}_s(\boldsymbol{p}')e^{-i\mathbf{p}'\cdot\mathbf{x}_1}, \tag{10.57}$$

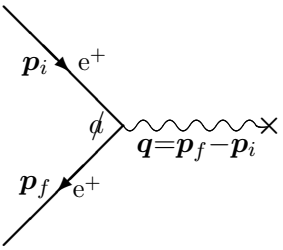


Figure 10.3: Scattering of positron by an external Coulomb field $\not{A}$. $\boldsymbol{p}_i$ and $\boldsymbol{p}_f$ denote the momenta of the incoming and scattered positron and $\boldsymbol{q}$, the three-momentum transfer to the positron.

we get

$$\mathcal{M} = \frac{ie}{\hbar c}\frac{1}{(2\pi)^3}\int d^4x_1 \int d^3p \int d^3p' \sum_{r,s}^{1,2} \langle 0|d_{s_f}(\boldsymbol{p}_f)d_r^\dagger(\boldsymbol{p})d_s(\boldsymbol{p}')d_{s_i}^\dagger(\boldsymbol{p}_i)|0\rangle \times \frac{1}{\sqrt{2E_p 2E_{p'}}}\bar{v}_s(\boldsymbol{p}')\not{A}^e(\mathbf{x}_1)v_r(\boldsymbol{p})e^{i(\mathbf{p}-\mathbf{p}')\cdot\mathbf{x}_1}. \quad (10.58)$$

It can be easily verified that

$$\langle 0|d_{s_f}(\boldsymbol{p}_f)d_r^\dagger(\boldsymbol{p})d_s(\boldsymbol{p}')d_{s_i}^\dagger(\boldsymbol{p}_i)|0\rangle = \delta_{rs_f}\delta_{ss_i}\delta(\boldsymbol{p}-\boldsymbol{p}_f)\delta_(\boldsymbol{p}'-\boldsymbol{p}_i). \quad (10.59)$$

Substituting (10.59) into Eq. (10.58) and integrating over d^3p and d^3p', we obtain

$$\mathcal{M} = \frac{ie}{\hbar c}\frac{1}{(2\pi)^3}\frac{1}{\sqrt{2E_{p_i}2E_{p_f}}}\int d^4x_1 \bar{v}_{s_i}(\boldsymbol{p}_i)\not{A}^e(\mathbf{x}_1)v_{s_f}(\boldsymbol{p}_f)e^{i(\mathbf{p}_f-\mathbf{p}_i)\cdot\mathbf{x}_1}. \quad (10.60)$$

Using the Fourier transform relation

$$\not{A}(\mathbf{q}) = \frac{1}{(2\pi)^4}\int \not{A}^e(\mathbf{x}_1)e^{i\mathbf{q}\cdot\mathbf{x}}d^4x,$$

we obtain the matrix element in the momentum representation.

$$\mathcal{M} = \left(\frac{ie}{\hbar c}\right)\frac{2\pi}{\sqrt{2E_{p_i}2E_{p_f}}}\bar{v}_{s_i}(\boldsymbol{p}_i)\not{A}(\mathbf{p}_f-\mathbf{p}_i)v_{s_f}(\boldsymbol{p}_f). \quad (10.61)$$

This is exactly the matrix element that one would have obtained in Feynman's theory. From the equation obtained for the matrix element,

one can frame the rule for drawing the Feynman diagram in momentum space for positron scattering. The initial and final states follow the path sequence. If we are considering positron scattering, the final state occurs at the bottom of the diagram (with time flowing upwards) as indicated in Fig. 10.3.

Thus we have established the connection between the field theoretical formalism and the much simpler Feynman formalism developed intuitively by him.

10.5.2 Second order term in S-matrix

Let us now consider the second order term in the S-matrix and show how helpful is Wick's theorem in expanding it into a sum of normal products which can be used to study a large variety of quantum electrodynamic processes. The second order term involves a product of two $\mathscr{H}_I(\mathbf{x})$ and the basic vertex diagrams which we obtained in the first order can be combined in a meaningful way to yield Feynman diagrams that represent real processes.

Using Wick's theorem, let us expand the second order term in the S-matrix as a sum of normal products.

$$\begin{aligned} & T\left(\mathscr{H}_I(\mathbf{x}_1)\mathscr{H}_I(\mathbf{x}_2)\right) \\ &= T\left\{N\left(\bar{\psi}(\mathbf{x}_1)\not{A}(\mathbf{x}_1)\psi(\mathbf{x}_1)\right) N\left(\bar{\psi}(\mathbf{x}_2)\not{A}(\mathbf{x}_2)\psi(\mathbf{x}_2)\right)\right\} \\ &= N_A + N_B + N_C + N_D + N_E + N_F + N_G + N_H, \end{aligned} \tag{10.62}$$

where

$$N_A = N\left(\bar{\psi}(\mathbf{x}_1)\not{A}(\mathbf{x}_1)\psi(\mathbf{x}_1)\bar{\psi}(\mathbf{x}_2)\not{A}(\mathbf{x}_2)\psi(\mathbf{x}_2)\right), \tag{10.63}$$

$$N_B = N\left(\bar{\psi}(\mathbf{x}_1)\not{A}(\mathbf{x}_1)\psi^{\bullet}(\mathbf{x}_1)\bar{\psi}^{\bullet}(\mathbf{x}_2)\not{A}(\mathbf{x}_2)\psi(\mathbf{x}_2)\right), \tag{10.64}$$

$$N_C = N\left(\bar{\psi}^{\bullet}(\mathbf{x}_1)\not{A}(\mathbf{x}_1)\psi(\mathbf{x}_1)\bar{\psi}(\mathbf{x}_2)\not{A}(\mathbf{x}_2)\psi^{\bullet}(\mathbf{x}_2)\right), \tag{10.65}$$

$$N_D = N\left(\bar{\psi}(\mathbf{x}_1)\gamma^{\alpha}A^{\bullet}_{\alpha}(\mathbf{x}_1)\psi(\mathbf{x}_1)\bar{\psi}(\mathbf{x}_2)\gamma^{\alpha}A^{\bullet}_{\alpha}(\mathbf{x}_2)\psi(\mathbf{x}_2)\right), \tag{10.66}$$

$$N_E = N\left(\bar{\psi}(\mathbf{x}_1)\gamma^{\alpha}A^{\bullet}_{\alpha}(\mathbf{x}_1)\psi^{\bullet\bullet}(\mathbf{x}_1)\bar{\psi}^{\bullet\bullet}(\mathbf{x}_2)\gamma^{\alpha}A^{\bullet}_{\alpha}(\mathbf{x}_2)\psi(\mathbf{x}_2)\right), \tag{10.67}$$

$$N_F = N\left(\bar{\psi}^{\bullet}(\mathbf{x}_1)\gamma^{\alpha}A^{\bullet\bullet}_{\alpha}(\mathbf{x}_1)\psi(\mathbf{x}_1)\bar{\psi}(\mathbf{x}_2)\gamma^{\alpha}A^{\bullet\bullet}_{\alpha}(\mathbf{x}_2)\psi^{\bullet}(\mathbf{x}_2)\right), \tag{10.68}$$

$$N_G = N\left(\bar{\psi}^{\bullet}(\mathbf{x}_1)\not{A}(\mathbf{x}_1)\psi^{\bullet\bullet}(\mathbf{x}_1)\bar{\psi}^{\bullet\bullet}(\mathbf{x}_2)\not{A}(\mathbf{x}_2)\psi^{\bullet}(\mathbf{x}_2)\right), \tag{10.69}$$

$$N_H = N\left(\bar{\psi}^{\bullet}(\mathbf{x}_1)\gamma^{\alpha}A^{\bullet\bullet}_{\alpha}(\mathbf{x}_1)\psi^{\bullet\bullet\bullet}(\mathbf{x}_1)\bar{\psi}^{\bullet\bullet\bullet}(\mathbf{x}_2)\gamma^{\alpha}A^{\bullet\bullet}_{\alpha}(\mathbf{x}_2)\psi^{\bullet}(\mathbf{x}_2)\right). \tag{10.70}$$

The first term N_A is not very interesting. It corresponds to two disconnected vertex diagrams and does not represent any real process.

The second and third terms N_B and N_C give identical results since permutation of Fermion operators in the two groups $(\bar{\psi}\not{A}\psi)_{\mathbf{x}_1}$ and $(\bar{\psi}\not{A}\psi)_{\mathbf{x}_2}$

is even and the space-time variables are dummy variables that are integrated. Expressions (10.64) and (10.65) contain one Fermion contraction which represents the Fermion propagator from one space-time point to another besides two uncontracted Fermion operators and two uncontracted photon operators. The uncontracted operators annihilate the particles that are present initially and emit the particles that are present finally. Since the operators are arranged in normal order, it is easy to pick out the terms that contribute to a given process. The terms N_B and N_C can describe several processes involving two Fermions and two photons such as Compton scattering by electron, Compton scattering by positron, electron-positron pair annihilation into two photons and electron-positron pair creation by two photons.

$$\begin{array}{llll} (i) & \gamma + e^- & \rightarrow & \gamma + e^-, \\ (iii) & e^+ + e^- & \rightarrow & \gamma + \gamma, \end{array} \qquad \begin{array}{llll} (ii) & \gamma + e^+ & \rightarrow & \gamma + e^+, \\ (iv) & \gamma + \gamma & \rightarrow & e^+ + e^-. \end{array}$$

The fourth term N_D consists of four uncontracted Fermion operators and one photon-photon contraction representing the photon propagation between the two Fermions. By picking out suitable Fermion annihilation and creation operators from the uncontracted Fermion operators, we can describe electron-electron scattering, electron-positron scattering and positron-positron scattering. For each of these processes, there will be two Feynman diagrams. For electron-electron and positron-positron scattering, there will be two Feynman amplitudes, one for direct scattering and another for exchange scattering with a relative change in sign. In non-relativistic quantum mechanics, the relative minus sign arises from the use of antisymmetric wave function of Fermions. In the field theory, this is provided by the anticommutative property of the Fermion field operators. In the case of electron-positron scattering (known as Bhabha scattering), there are also two Feynman diagrams, one direct scattering and the other representing electron-positron pair annihilation into a photon and subsequent pair creation. These two modes are indistinguishable and should be taken into account in the calculation.

The fifth and sixth terms N_E and N_F can be clubbed together. Each term consists of contraction of one Fermion pair and one photon-pair, yielding one Fermion propagator and one photon propagator. The electron emits a photon which is recaptured by the same electron. This is known as *the self-energy of the electron* and is represented by the Feynman diagram 10.4(a). This process converts the *bare* electron into a *physical* electron. The self-interaction energy changes the mass of the bare elec-

tron into that of the physical electron. Its evaluation leads to a divergent integral which is overcome by the method of *renormalization.*

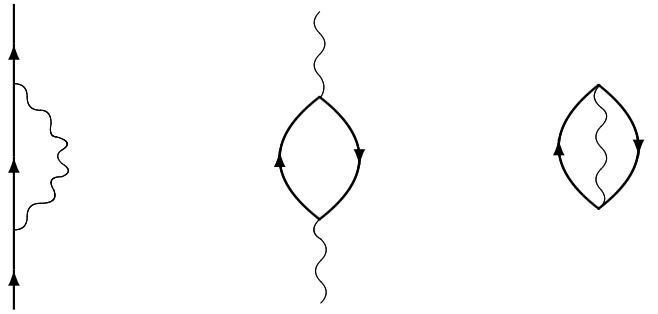


Figure 10.4: Feynman diagrams representing (a) Self energy of the electron (b) Self energy of the photon or Vacuum polarization and (c) Vacuum diagram.

The seventh term N_G consists of two Fermion pair contractions and two uncontracted photon field operators. The photon creates a virtual electron-positron pair which again annihilates into a photon. This is known as the *self-energy of the photon* and is depicted by the Feynman diagram (b) in Fig. 10. This is also known as *vacuum polarization* since the external electromagnetic field such as that of heavy nucleus can modify the distribution of virtual electron-positron pairs in the vacuum. The calculation of the photon self-energy also leads to infinities which are again eliminated by renormalization procedure.

The last term N_H is one in which the two Fermion-pairs and one photon-pair are contracted leaving no external line. This is a vacuum diagram as illustrated in Fig. 10.4(c). Such *vacuum diagrams* without any external line can be omitted.

Review Questions

10.1 Develop an invariant perturbation theory treating the interaction between fields as a small perturbation and deduce the S-matrix formalism for the study of scattering phenomena.

10.2 Define Normal Products, Dyson's Chronological products, Wick's chronological products and Wick's contraction. State Wick's theorem and explain how it is useful in reducing Wick's chronological product into normal products.

10.3 Consider the first order term in the S-matrix expansion for the interaction of the Dirac field with an external electromagnetic field and explain how you can represent the various processes by means of Feynman diagrams. Deduce the Feynman matrix element in momentum representation for the positron scattering by an external electromagnetic field.

10.4 Explain how the S-matrix can be decomposed into normal products using Wick's theorem. Illustrate your answer by considering the second order term in the S-matrix for the electro-magnetic interaction.

Problems

10.1 Given Eq. (10.3)

$$(H_0 + H_i)\Phi(t) = i\hbar\frac{\partial\Phi(t)}{\partial t},$$

where H_0 and H_i are the unperturbed and interaction Hamiltonian and $\Phi(t)$ is the time-dependent state vector in the Schrödinger picture, deduce Eq. (10.5)

$$i\hbar\frac{\partial\Psi(t)}{\partial t} = H_I(t)\Psi(t),$$

in the interaction picture where the state vector $\Psi(t)$ and the interaction Hamiltonian $H_I(t)$ are defined by

$$\Psi(t) = e^{iH_0t/\hbar}\Phi(t) \quad \text{and} \quad H_I(t) = e^{iH_0t/\hbar}H_ie^{-iH_0t/\hbar}.$$

10.2 Given the Dirac spinors and the gamma matrices, construct a scalar, a pseudoscalar, a four-vector, a pseudo vector, and an antisymmetric tensor of second rank. With the help of these, construct an interaction Hamiltonian for electron-photon interaction and nucleon-pion interaction.

10.3 It is known that parity is not conserved in weak interaction. So, construct an interaction Hamiltonian for the parity-violating β-decay.

Solutions to Problems

10.1 Since $\Psi(t) = e^{iH_0t/\hbar}\Phi(t)$ and $i\hbar\frac{\partial\Phi(t)}{\partial t} = (H_0 + H_i)\Phi(t)$, we get by differentiating Ψ,

$$\begin{aligned}
\frac{\partial\Psi}{\partial t} &= \frac{i}{\hbar}H_0e^{iH_0t/\hbar}\Phi(t) + e^{iH_0t/\hbar}\frac{\partial\Phi}{\partial t} \\
&= \frac{i}{\hbar}H_0e^{iH_0t/\hbar}\Phi(t) + e^{iH_0t/\hbar}\frac{1}{i\hbar}(H_0 + H_i)\Phi(t) \\
&= \frac{1}{i\hbar}e^{iH_0t/\hbar}H_i\Phi(t) \\
&= \frac{1}{i\hbar}e^{iH_0t/\hbar}H_ie^{-iH_0t/\hbar}\Psi(t).
\end{aligned}$$

Hence it follows that

$$i\hbar\frac{\partial\Psi}{\partial t} = H_I(t)\Psi(t),$$

where $H_I(t) = e^{iH_0t/\hbar}H_ie^{-iH_0t/\hbar}$.

10.2 Given the Dirac spinors and γ matrices, we can construct the following quantities which have definite transformation property:

Quantity	Transformation Property	No. of components
$\bar{\psi}\psi$	Scalar	1
$\bar{\psi}\gamma_5\psi$	Pseudoscalar	1
$\bar{\psi}\gamma^\mu\psi$	Four-vector	4
$i\bar{\psi}\gamma_\mu\gamma_5\psi$	Pseudo-vector	4
$\frac{i}{2}\bar{\psi}(\gamma^\mu\gamma^\nu - \gamma^\nu\gamma^\mu)\psi$	Antisymmetric tensor (2^{nd} rank)	6

The factor i is included in order to take care of the Hermitian property.

Since the photon field is represented by a four-vector A_μ, the electron-photon interaction is represented by the interaction Hamiltonian

$$H_{\text{int}} = \bar{\psi}\gamma^\mu\psi A_\mu,$$

which is a scalar under Lorentz transformation.

The pion-nucleon interaction is a bit complicated. Since pion has an intrinsic negative parity when compared to the nucleon, the interaction Hamiltonian will have a form

$$\bar{\psi}_N\gamma_5\psi_N\phi,$$

where ϕ represents the pseudo-scalar pion field. This interaction Hamiltonian is a scalar under both Lorentz transformation and parity operation. Since the pion exists in three charge states, it can be represented by an isospin vector $\boldsymbol{\tau}_\pi$. Since the nucleon exists in two charge states, it can be represented by a vector $\boldsymbol{\tau}_N$ in isospin space. The interaction Hamiltonian has to be a scalar in the iso-spin space too. This yield an interaction Hamiltonian

$$\bar{\psi}_N\gamma_5\boldsymbol{\tau}_N\cdot\boldsymbol{\tau}_\pi\psi_N\phi_\pi,$$

where ψ_N is a two-component vector and ϕ_π is a three-component vector in the isospin space.

$$\psi_N = \begin{bmatrix} \psi_p \\ \psi_n \end{bmatrix}; \qquad \phi_\pi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix}.$$

10.3 It is known that the beta decay interaction is a combination of vector and axial vector interaction.

$$H_{\text{int}} = G_\beta \sum_{i=V,A} C_i(\bar{\psi}_p \mathcal{O}_i \psi_n)(\bar{\psi}_e \mathcal{O}_i \psi_\nu),$$

where G_β is the weak interaction coupling constant and C_V and C_A are the coupling coefficients for vector and axial vector interaction. The operators $\mathcal{O}_V = \gamma^\mu$ and $\mathcal{O}_A = \gamma^\mu\gamma_5$ This is a parity conserving interaction. Subsequently Lee and Yang postulated that weak interaction need not be conserved in weak interaction and it has been experimentally confirmed by Wu *et al.* It is found that $C_A = -C_V$ for pure leptonic weak interaction such as μ-decay. Incorporating this in $V - A$ interaction, we get

$$H_{\text{int}} = \frac{G_\beta}{\sqrt{2}} C_V[\bar{\psi}_p \gamma^\mu (1 - \gamma_5)\psi_n][\bar{\psi}_e \gamma_\mu (1 - \gamma_5)\psi_\nu].$$

The inclusion of the parity non-conserving extra term requires the insertion of a factor $1/\sqrt{2}$ in order to retain the old definition of the coupling constant G_β.

Chapter 11

Recent Developments

The purpose of the present book is limited to the study of Relativistic Quantum Mechanics and its extension to Feynman's formulation of Quantum Electrodynamics and highlight how Feynman's formulation is equivalent to the more general approach of Quantum Field Theory, the elements of which have been discussed in some detail. But, it is found desirable to give a bird's eye view of the difficulties encountered in treating higher order terms in the perturbation theory and the recent advances in Quantum Field Theory. The divergent integrals that one meets in the perturbation theory have been overcome by evolving renormalization techniques. It is shown that Quantum Electrodynamics is a gauge theory that is renormalizable. Recent advances in Quantum Field Theory include the successful attempts to bring in its fold the weak interactions and strong interactions. So, the Quantum Field Theory has emerged as a fundamental theory of electromagnetic, weak and strong interactions, leading to the formulation of the Standard Model of elementary particles. It is being generally felt that the Quantum Field Theory is a low energy limit of a more fundamental theory which includes gravitational interaction. The String Theory and Superstring Theory appear to be promising candidates for such a general theory of everything.

11.1 The Renormalization Program

As long as we are considering the QED processes in the lowest order of perturbation expansion, we face no problems and we get results that tally

fairly well with the experimental data. If we wish to improve the accuracy of our calculations and go to higher orders, we encounter divergent integrals as explained in Ch. 5. The integral[1]

$$\int \frac{d^4k}{(2\pi)^4} \frac{1}{\mathbf{k}^4}$$

diverges logarithmically. Many have struggled with these infinities during 1940s. Eventually, they have found a way out by using different techniques of regularization and renormalization. By renormalizing mass, charge and wavefunctions, one can get rid of infinities in all orders of perturbation theory. A field theory which gives finite answers for all observable phenomena to all orders in perturbation theory by renormalization procedure is said to be *renormalizable*. QED is indeed a renormalizable theory.

11.1.1 Regularization and Renormalization

In evaluating the integral $\int \frac{d^4k}{(2\pi)^4}$, we simply integrate only upto Λ, known as a cut off, to avoid the divergences. Then the integral is said to be regularized. After regularization, we speak of cut-off dependent quantities instead of divergent quantities.

But any physically observable quantity such as scattering cross section should not depend upon Λ.

In the S-matrix formalism, when the interaction is switched off, we deal with only bare particles without self-interaction *i.e.* with only bare mass m_0 and bare charge e_0. But these bare particles are not observable and all the observable attributes are experimentally observable mass m and charge e. The difference between the bare mass m_0 and experimental mass m and similarly the difference between bare charge e_0 and the observable charge e depend upon the cut-off parameter Λ but the physically observable quantities such as the scattering cross section depend only upon quantities such as experimental mass and charge. This is what we call renormalization. The physically observable quantities are indeed independent of the cut-off parameter. There are two different ways of achieving this. One is called the *Pauli-Villars regularization* and the other is known as the *dimensional regularization.*

[1]The integral $\int^\infty dr\, r^n$ diverges linearly for $n = 0$, quadratically for $n = 1$ and so on and $\int^\infty \frac{dr}{r}$ diverges logarithmically.

11.2 Gauge Theories

11.2.1 Gauge theory of electromagnetic interaction

Consider the Lagrangian of a complex scalar field.

$$\mathscr{L}_\phi = (\partial^\mu \phi^\dagger)(\partial_\mu \phi) - m^2 \phi^\dagger \phi. \tag{11.1}$$

It is invariant under a global phase rotation which is independent of the local space-time coordinates, since each term in the Lagrangian consists of ϕ and $\phi^\dagger$.

$$\phi \longrightarrow \phi' = \exp(-ieQ\theta)\phi, \tag{11.2}$$

where e is the unit of positive electric charge and Q is the charge of the scalar particle in that unit and θ is the global parameter which is independent of space-time coordinates.

In a similar way, let us consider the invariance of the Lagrangian of the Dirac field under such a global phase rotation.

$$\mathscr{L}_\psi = \bar{\psi}(i\partial\!\!\!/ - m)\psi, \tag{11.3}$$

with $\partial\!\!\!/ = \gamma^\mu \partial_\mu = \gamma^\mu \frac{\partial}{\partial x^\mu}$. The invariance of the action under a global transformation

$$\psi \to \psi' = e^{-ieQ\theta}\psi, \tag{11.4}$$

implies a conserved current j^μ by Noether's theorem.

$$j^\mu = eQ\bar{\psi}\gamma^\mu\psi \tag{11.5}$$

The corresponding conserved charge is

$$\begin{aligned} q &= \int d^3x\, j^0 = eQ \int d^3x \bar{\psi}\gamma^0\psi = eQ \int d^3x \psi^\dagger\psi \\ &= eQ \int d^3p \sum_r \left(c_r^\dagger(\boldsymbol{p}) c_r(\boldsymbol{p}) - d_r^\dagger(\boldsymbol{p}) d_r(\boldsymbol{p}) \right), \end{aligned} \tag{11.6}$$

where $c_r, c_r^\dagger, d_r$ and $d_r^\dagger$ are the annihilation and creation operators of the Fermions and anti-Fermions. The concept of conserved charge can be extended to include the conservation of lepton number or baryon number,

since Eq. (11.6) refers to any property that depends on the number of Fermions minus the number of anti-Fermions.

Let us investigate what happens if we impose the global phase rotation symmetry to be also a local symmetry. In this case, θ is a function of space-time coordinates x^μ. From Eq. (11.4), it follows

$$\begin{aligned} \partial_\mu \psi' &= \partial_\mu e^{-ieQ\theta}\psi \\ &= (\partial_\mu e^{-ieQ\theta})\psi + e^{-ieQ\theta}\partial_\mu\psi \\ &= -ieQ(\partial_\mu\theta)e^{-ieQ\theta}\psi + e^{-ieQ\theta}\partial_\mu\psi. \end{aligned} \tag{11.7}$$

For Fermions,

$$\begin{aligned} \mathscr{L}_{\psi'} &= \bar{\psi}'(i\partial\!\!\!/ - m)\psi' \\ &= \bar{\psi}eQ(\partial_\mu\theta)\gamma^\mu\psi + \bar{\psi}(i\partial\!\!\!/ - m)\psi \\ &= eQ(\partial_\mu\theta)\bar{\psi}\gamma^\mu\psi + \mathscr{L}_\psi. \end{aligned} \tag{11.8}$$

Equation (11.8) clearly indicates that the Dirac Lagrangian is not invariant under local phase rotation. This failure can be rectified by introducing new fields such that the modified Lagrangian has the local symmetry. The extra term on the right-hand side of Eq. (11.8) involves a factor $\partial_\mu\theta$ which transforms like a four-vector. Introducing an additional term with a four-vector field A_μ in the Dirac Lagrangian, a modified Lagrangian

$$\mathscr{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi - eQ\bar{\psi}\gamma^\mu\psi A_\mu, \tag{11.9}$$

can be obtained which has a local symmetry under the transformation

$$\psi \to \psi' = e^{-ieQ\theta(\mathbf{x})}\psi, \qquad A_\mu \to A'_\mu. \tag{11.10}$$

What should be the transformation property of A_μ such that the modified Lagrangian is invariant? The modified Lagrangian $\mathscr{L}'$

$$\begin{aligned} \mathscr{L}' &= \bar{\psi}'(i\partial\!\!\!/ - m)\psi' - eQ\bar{\psi}'\gamma^\mu\psi' A'_\mu \\ &= \bar{\psi}(i\partial\!\!\!/ - m)\psi - eQ(A'_\mu - \partial_\mu\theta)\bar{\psi}\gamma^\mu\psi, \end{aligned} \tag{11.11}$$

is invariant if the newly introduced four-vector field A_μ transforms as

$$A'_\mu = A_\mu + \partial_\mu\theta. \tag{11.12}$$

Now one can easily identify that the newly introduced field A_μ is none other than the electromagnetic field which leaves the Lagrangian of the

free electromagnetic field invariant under the same local transformation. Adding the free field Lagrangian of the electromagnetic field, we obtain the full Lagrangian for the combined field of ψ and A_μ.

$$\mathscr{L} = \bar{\psi}(i\partial\!\!\!/ - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - eQ\bar{\psi}\gamma^\mu\psi A_\mu. \tag{11.13}$$

This is the Lagrangian that is used in field theory for the study of Quantum Electrodynamics (QED). It may be observed that the full Lagrangian (11.13) consists of an interaction term

$$\mathscr{L}_{\text{int}} = -eQ\bar{\psi}\gamma^\mu\psi A_\mu,$$

besides the free-field Lagrangians of Dirac and e.m. fields. Since the origin of the interaction term can be traced to the introduction of a new field to make the Dirac Lagrangian invariant under a local gauge transformation, the newly introduced field has come to be known as *gauge field*, its field quantum as *gauge Boson* and the underlying theory as *gauge theory*.

The same technique can be used for any other field with global symmetry. Let us illustrate it for the charged scalar field ϕ. For this, let us define a *gauge covariant derivative*[2] D_μ.

$$D_\mu = \partial_\mu + ieQA_\mu, \tag{11.14}$$

such that the complete Lagrangian for QED (Eq. (11.13)) can be written in the form

$$\mathscr{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{11.15}$$

Any complex scalar field ϕ which has a global phase symmetry for the transformation

$$\phi \to e^{-ieQ\theta}\phi$$

will also acquire a local phase symmetry with the substitution

$$\partial_\mu \to D_\mu; \qquad A_\mu \to A_\mu + \partial_\mu\theta.$$

Thus we can obtain the Lagrangian for the scalar field interacting with the electromagnetic field as

$$\mathscr{L} = (D_\mu\phi)^\dagger(D_\mu\phi) - m^2\phi^\dagger\phi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \tag{11.16}$$

[2]The gauge covariant derivative D_μ transforms in the same way as ψ.

$$\psi(\mathbf{x}) \to e^{-ieQ\theta(\mathbf{x})}\psi(\mathbf{x}); \qquad D_\mu\psi(\mathbf{x}) \to e^{-ieQ\theta(\mathbf{x})}D_\mu\psi(\mathbf{x}).$$

11.2.2 Spontaneous symmetry breaking

The gauge Boson that mediates the electromagnetic interaction is massless, whereas the Bosons $W^{\pm}, Z^0$ that mediate the weak interaction are massive. The problem now is how to give mass to the weak gauge Bosons without breaking the gauge symmetry of the Lagrangian, which is essential for a renormalizable theory.

Consider a simple example of a U(1) gauge theory, describing electromagnetic interaction of a charged scalar (spin 0) field ϕ. The corresponding Lagrangian is

$$\begin{aligned} \mathscr{L}_{EM} &= (\partial^\mu - ieA^\mu)\phi^*(\partial_\mu + ieA_\mu)\phi - [\mu^2\phi^*\phi + \lambda(\phi^*\phi)^2] \\ &\quad -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \end{aligned} \tag{11.17}$$

where A^μ is the electromagnetic field and $F^{\mu\nu}$, the corresponding field tensor.

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu.$$

Since ∂_μ is the momentum operator, the first term represents the kinetic energy (K.E.) of the scalar field ϕ and also the gauge interaction. The last term in Eq, 11.17) represents the gauge kinetic energy. The middle term represents the scalar mass and self interaction and it is invariant under the phase transformation

$$\phi \rightarrow e^{i\alpha(\mathbf{x})}\phi, \tag{11.18}$$

which is usually called the gauge transformation. But $\partial_\mu\phi$ and the resulting scalar K.E. are not invariant under the phase (gauge) transformation. However, as we saw earlier, the invariance of the Lagrangian under the gauge transformation on ϕ is restored if one makes a simultaneous transformation on A_μ as given below:

$$A_\mu \rightarrow A_\mu = \frac{1}{e}\partial_\mu\alpha(\mathbf{x}). \tag{11.19}$$

Thus it is a remarkable property of gauge theory that the symmetry of the Lagrangian under the phase transformation of a charged particle field implies the interaction dynamics.

The symmetry of the Lagrangian will be lost if we add a mass term $M^2A^\mu A_\mu$, which is clearly not invariant under the simultaneous transformations of (11.18) and (11.19). While the photon has no mass, the

analogous gauge Bosons $W^{\pm}, Z^0$ for weak interaction are massive. Now the question arises how to add such mass terms without destroying the gauge symmetry of the Lagrangian.

The clue comes from the observation that the scalar mass term $\mu^2\phi^*\phi$ is gauge invariant. This is exploited to give mass to $W^{\pm}, Z^0$ surreptitiously by allowing them to acquire mass by absorbing scalar particles. This is similar to the situation of "a snake acquiring mass by swallowing a rabbit". This is the famous *Higgs mechanism*.

Let us start with a scalar field of mass μ^2. The corresponding scalar potential is

$$V(\phi) = \mu^2\phi^*\phi + \lambda(\phi^*\phi)^2, \tag{11.20}$$

which is depicted in Fig. 11.1

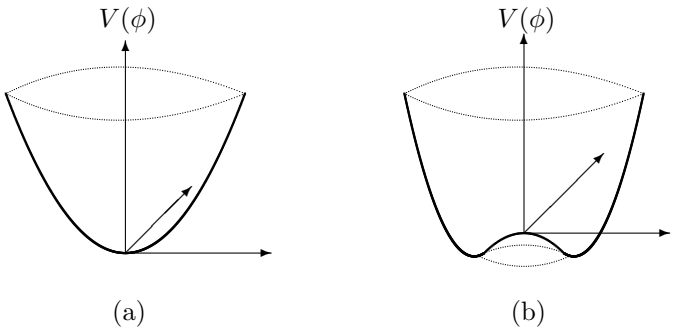


Figure 11.1: The shape of $V(\phi)$ for (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$

The potential $V(\phi)$ will not have any minimum if λ is negative, since the potential will continue to decrease with increase of $|\phi|$. So, we conclude that λ has to be positive. If μ^2 is positive, then the minimum value $V(\phi) = 0$ occurs at $\phi = 0$ and $V(\phi)$ increases with ϕ, as shown in Fig. 11.1(a). The more interesting case corresponds to negative μ^2, as illustrated in Fig. 11.1(b). The potential $V(\phi)$ is minimum at the value of ϕ that corresponds to

$$\frac{dV}{d\phi} = 0; \qquad \phi = v = \sqrt{-\frac{\mu^2}{2\lambda}}. \tag{11.21}$$

Thus the ground state (vacuum state) corresponds to this finite value of $|\phi|$.

Since perturbation expansion should be done around the point of minimum energy (stable equilibrium), we must translate the origin to this point and perform perturbation with the redefined field

$$h(\mathbf{x}) = \phi(\mathbf{x}) - v, \tag{11.22}$$

which corresponds to a physical scalar field, called *Higgs Boson*. Substituting (11.22) into the Lagrangian (11.17), we find several surprises. The first term of the Lagrangian gives a quadratic term in A^μ

$$e^2 v^2 A^\mu A_\mu,$$

which corresponds to a gauge Boson of mass

$$M = ev.$$

Moreover, the middle term leads to a real mass for the physical scalar.

$$M_h = \sqrt{-\mu^2} = (\sqrt{2\lambda})v = \frac{\sqrt{2\lambda}}{e} M. \tag{11.23}$$

Thus the mass of Higgs Boson is equal to the mass of the gauge Boson, multiplied by a quantity which is the ratio of the self coupling to the gauge coupling. Although we do not know the value of the self-coupling, the validity of perturbation theory implies that $\lambda \leq 1$. This means that the mass of Higgs Boson is expected to be of the same order as the mass of the weak interaction Boson. Hence it is a prime candidate for discovery in Large Hadron Collider (LHC) experiments.

Once again, let us look at Fig. 11.1 to appreciate the *Higgs mechanism*. The choice of negative μ^2 meant that the ground state has moved from the origin to a finite value of $|\phi|$ which corresponds to any point on the rim. Thus the phase (gauge) symmetry of the Lagrangian is not shared by the ground state (vacuum state). This phenomenon is known as *spontaneous symmetry breaking*.

The breaking of the phase (gauge) symmetry by the ground state enables the gauge boson to acquire mass. At the same time, the fact that this symmetry is retained by the Lagrangian enables the renormalization to go through. This last feature is found to be very important because it ensures

a systematic cancellation of infinities, without which there will not be any predictable field theory. The renormalizability of electro-weak gauge theory in the presence of spontaneous symmetry breaking was shown to be valid by t'Hooft[3] for which he was awarded the Nobel Prize in 1999.

Let us now summarize the results of our study. The requirement of invariance under a local gauge transformation is essential for finding the equations describing fundamental interaction. Take for example the Lagrangian for a free electron. The requirement that the Lagrangian should be locally invariant under the same type of transformation can only be fulfilled by introducing additional terms. The form of these terms is determined by the symmetry requirement, which results in the introduction of the electromagnetic field. In a sense, the electromagnetic field is a consequence of the local symmetry of the Lagrangian for the electron.

This procedure can be generalized to more complex transformations (for example referring to mixing the components of field operators) and new interactions. By requiring local gauge invariance additional fields can be introduced. These additional fields describe the interaction between the original fields. The gauge principle provides a general scheme for introducing interaction by constructing gauge field theories. To this end one starts with a Lagrangian for a matter field and derives the interaction by introducing exactly those fields that make the Lagrangian invariant under a relevant local gauge transformation. It seems that all fundamental forces can be described by such local gauge field theories. Gauge symmetry plays a crucial role in determining the dynamics of the theory since the nature of gauge transformation determines the possible interaction. The structure of these transformations are characterized by special mathematical groups: U(1) for QED, SU(2)×U(1) for electroweak interaction, SU(3) for strong interaction. The relations between these groups are exploited in programs for the unification of the fundamental types of interaction. From a more technical point of view, gauge symmetries are important tools in proofs of renormalizability. So, it turns out that gauge invariance plays as much an important role as Lorentz invariance in the selection of theories. Since gauge invariance plays a pivotal role in the discovery of quantum field theories, it is a paradigm case for how a rich mathematical structure can help in the construction of theories.

[3] t Hooft, G. (1971). "Renormalization of massless Yang-Mills fields". Nuclear Physics **B 33**, 173-177 (1971); "Renormalizable Lagrangians for massive Yang-Mills fields". Nuclear Physics **B 35** 167-448 (1971).

11.3 The Standard Model

It has been established that there are four different types of forces that operate in this universe: (1) Strong interaction (2) Electro-magnetic interaction (3) Weak interaction and (4) Gravitational interaction. It is the ultimate aim of the physicists to formulate a unified theory of all these four forces into one integrated "theory of everything".

Of the four interactions, electromagnetic interaction has been fairly well understood and all the processes in Quantum Electrodynamics can be explained to a high degree of accuracy on the basis of the theory developed by Tomonago, Schwinger and Feynman. The electro-magnetic interaction is mediated by photons (gauge Bosons) and in the field theoretic parlance, it is a gauge theory with gauge group U(1). The weak interaction is mediated by intermediate vector bosons W^+, W^-, Z^0 and Sheldon Glashow, Steven Weinberg and Abdus Salam developed a unified theory of electromagnetic and weak interactions which has come to be known as electro-weak theory which is a gauge theory of gauge group SU(2) $\times$ U(1). The electro-weak theory has been enlarged to include strong interaction which arises from exchange of gluons between the coloured quarks that constitute the hadrons. The extended theory unifies the electro-weak interactions with strong interactions and it is popularly known as the standard model of particle physics. It is a gauge theory of the electroweak and strong interactions with the gauge group SU(3) $\times$ SU(2) $\times$ U(1).

11.3.1 Basic elementary particles - Fermions

The Standard Model is built on the basis of 12 elementary particles of spin 1/2 known as Fermions which obey the Pauli exclusion principle. Each Fermion has a corresponding anti-particle. They are classified into quarks and leptons, depending upon their interactions. There are six quarks (up, down, strange, charm, bottom, top) and six leptons (electron, electron neutrino, muon, muon neutrino, tau, tau neutrino). Pairs of these particles from each group form a set, known as a generation, as shown in Table. 11.1 There are three generations.

The quarks carry colour charge and they interact by exchange of gluons which are also coloured objects. This results in the theory of colour forces known as quantum chromodynamics (QCD) and it has been devel-

Table 11.1: Fermions that form the basis for the Standard Model.

Fermions	Charge	Generation I		Generation II		Generation III	
Quarks	$+\frac{2}{3}$	Up	u	Charm	c	Top	t
	$-\frac{1}{3}$	Down	d	Strange	s	Bottom	b
Leptons	-1	Electron	e^-	Muon	μ^-	Tau	τ^-
	0	e-neutrino	ν_e	μ-neutrino	ν_μ	τ-neutrino	ν_τ

oped as the theory of strong interactions in analogy with QED. But the fundamental difference is that photons do not carry electric charges in QED but in QCD the gluons carry colour charges. QCD is a non-abelian gauge field theory of strong interactions and it has two important features:

1. At short distances *i.e.* at high momentum transfers, the interaction becomes weaker. This is known as *the asymptotic freedom.* In this region, one can resort to perturbation theory.

2. At large distances, the interaction grows stronger and stronger, leading to permanent confinement of quarks within colour-neutral composite particles (hadrons) consisting of either three quarks (baryons and hyperons) or either a quark-antiquark pair (mesons). This is known as *the infra-red slavery.*

Quarks also carry electric charge and weak isospin and hence they interact with other Fermions both by electromagnetic and weak interactions.

The remaining six Fermions do not carry colour charge and are called leptons. The three leptons, electron, muon and tau, are charged particles and they can interact through both electro-magnetic and weak interactions. The three neutrinos do not carry electric charge and hence they can interact only through weak interactions. This makes the detection of neutrinos an extremely difficult job.

Each member of a generation has greater mass than the corresponding particles of lower generations. The first generation charged particles do not decay. The second and third generation charged particles decay with very short half lives and are observed only in very high energy experiments. Neutrinos of all generations do not decay but pervade the entire universe without much interaction with baryonic matter.

11.3.2 Force mediating particles - Bosons

The standard model explains the interaction between particles to arise from an exchange of Bosons which are the force-mediating particles.

1. Photons mediate the electromagnetic interactions between charged particles. The photon is massless and the electromagnetic interaction is well-described by the theory of quantum electrodynamics.

2. The intermediate vector bosons W^+, W^-, Z^0 mediate the weak interaction between leptons and quarks of different flavours. They are massive, Z^0 being more massive than $W^\pm$. The weak interaction involving $W^\pm$ act exclusively on left-handed particles and right-handed antiparticles. Further, $W^\pm$ carry an electric charge of ± 1 and they couple to the electromagnetic interaction. The electrically neutral Z boson interacts with both left-handed particles and antiparticles. These three gauge bosons along with photon are clubbed together and they collectively mediate the electro-weak interactions.

3. The strong interaction between coloured quarks are mediated by eight gluons. The gluons are massless but they carry a colour which is labeled by a combination of colour and anti-colour charge (e.g. red-antigreen). Since the gluons have an effective colour charge, they can interact among themselves. The gluons and their interactions with themselves and with quarks are described by the theory of quantum chromodynamics.

11.3.3 The Higgs Boson

The Higgs particle is a massive scalar particle predicted by the Standard Model. It has no intrinsic spin and hence it is a Boson. The Higgs Boson plays a unique role in explaining how the elementary particles other than the photon and gluons acquire masses. It explains why the photon has no

mass while the W and Z Bosons are very heavy. In electro-weak theory, the Higgs Boson generates the masses of leptons (electron, muon and tau particle) and quarks. It is the only fundamental particle predicted by the Standard Model that has yet to be observed. There are already some indirect evidences and it is hoped that the Large Hadron Collider at CERN will confirm the existence of this particle.

11.3.4 Theoretical framework

The quantum field theory provides the theoretical framework for the Standard Model. Each particle is described in terms of a dynamical field that pervades the entire space-time. We first postulate a set of symmetries for the system and then write down the most general renormalizable Lagrangian for the system that observes these symmetries.

The global Poincaré symmetry is postulated for all relativistic quantum field theories. This symmetry includes the familiar translational symmetry, rotational symmetry and invariance under Lorentz transformation. The local SU(3)×SU(2)×U(1) gauge symmetry is an internal symmetry that essentially defines the standard model. The fields fall into different representations of the various symmetry groups of the Standard Model. The most general Lagrangian that displays the dynamics of the system involves 19 parameters whose numerical values are established by experiments.

Besides the symmetries postulated in the construction, the Standard Model exhibits four additional global symmetries, which are collectively denoted as *accidental symmetries*. These are continuous U(1) global symmetries which leave the Lagrangian invariant.

By Noether's theorem, each symmetry has an associated conservation law: the conservation of baryon number, electron number, muon number and tauon number. Each quark is assigned a baryon number 1/3 while each antiquark is assigned a baryon number $-1/3$. Conservation of baryon number implies that the number of quarks minus the number of antiquarks is a constant. This is strongly supported by experiments and no violation has been observed for this conservation law.

Similarly, each electron and its associated neutrino is assigned an electron number $+1$ whereas each of its antiparticle is assigned an electron number -1. In the same way, μ^- and ν_μ are assigned a muon number $+1$ whereas μ^+ and $\bar{\nu}_\mu$ are assigned a muon number -1; τ^- and ν_τ are as-

signed a tauon number +1 whereas τ^+ and $\bar{\nu}_\tau$ are assigned a tuaon number −1. The standard model predicts that each of these lepton family numbers should be conserved separately in a way similar to the conservation of baryon number.

The standard model predicts that neutrinos are massless. But experiments indicate that neutrinos may have small masses, since neutrinos are found to oscillate between flavours, signaling that the conservation of lepton family number is likely to be violated.

In addition to the accidental symmetries described above, the standard model exhibits several approximate symmetries.

11.3.5 Successes and failures of the Standard Model

The Standard Model predicted the existence of W and Z bosons, gluons, and the charm and top quarks before they were observed. Their predicted properties were experimentally confirmed with great precision as shown in Table. 11.2. The Standard Model also made several predictions about the decay of Z bosons, which were confirmed by the Large Electron-Positron Collider experiments at CERN.

Table 11.2: Predictions of the masses of W and Z bosons by the Standard Model and their experimentally measured values.

Quantity	SM prediction (GeV)	Experiment (GeV)
Mass of W Boson	80.390 ± 0.018	80.398 ± 0.025
Mass of Z Boson	91.1874 ± 0.0021	91.1876 ± 0.0021

There is some experimental evidence for the oscillations of neutrino from one flavour to the other. This will be possible only if the neutrinos possess some mass as against the zero mass prediction by the standard model. If the neutrinos have small masses, then there will be some violation of the lepton number flavour conservation.

One of the predictions of the Standard Model is the Higgs Boson which is yet to be experimentally observed. The Standard Model involves 19 numerical constants whose values are unrelated and arbitrary. and it does not allow finite, although small, masses for neutrinos.

Gravitation is out of the purview of the Standard Model and there is

no way of including the general relativity. So, the Standard Model does not explain the initial conditions of the universe that give rise to certain observed properties of the present-day universe, properties such as the predominance of matter over antimatter (matter/antimatter asymmetry), and its isotropy and homogeneity over large distances.

11.4 Gravitational Interaction

We have seen that the Standard Model explains fairly well the world of elementary particles, on the basis of electromagnetic, weak and strong interactions. The only other force left out is gravitational interaction. That means the gravitational interaction is so weak compared to electromagnetic, weak and strong interactions at the present day energies that the accelerators can reach. To get a feeling of the relative strength, consider the ratio of gravitational force to the electromagnetic force that operates between two protons:

$$\frac{F_{\text{gravity}}}{F_{\text{e.m.}}} = \frac{GM_p^2/r^2}{e^2/r^2} = 10^{-36} << 1, \tag{11.24}$$

where G is the gravitational constant and m_p is the rest mass of the proton. At high energies, one has to replace the rest mass of the proton by the effective relativistic mass

$$m_{\text{eff}} = \frac{m_p}{(1 - \frac{v^2}{c^2})^{1/2}}. \tag{11.25}$$

At high energies, when $m_{\text{eff}} = 10^{18} m_p$, the gravitational force is comparable to electromagnetic force and hence cannot be ignored. At such high energies, the Standard Model should fail.

From the general theory of relativity, it is conjectured that the gravitational field arises from an exchange of gravitons which are massless but have spin 2. This has motivated physicists to look for a more fundamental theory which will unify all the four forces that operate between the elementary particles at high energies that will reduce to the Standard Model at low energies.

Einstein's general theory of relativity relates the gravitational forces to the curved nature of space-time. This interlinking of gravity with space-time offers insurmountable difficulties in finding a quantum theory

of gravity that will unify all the four fundamental forces. Such a quantum theory should describe our universe, the physics of black holes and the processes involving elementary particles. String theory appears to be a promising candidate for such a unified theory.

11.5 Birth of String Theory

In the late 1960s, a totally novel idea emerged that particles should be treated as string-like objects instead of point-like objects under the name "dual models". This was to explain the experimentally observed behaviour of strongly interacting particles, called hadrons, at large angular momentum J. The experimental data on resonances (excited states of the light hadrons) of mass M and spin J fitted well with the relation $M^2 = J\alpha'$, where α' is called the Regge slope. These strings have infinitesimal thickness and tension $T = 1/(2\pi\alpha')$ and can be either open or closed. The particles can be identified as the various vibrational modes of an open string or closed string[4]. Since the vibrational modes predicted massless particles of spin 2 and higher spins which are not seen in the hadronic world, the theory was discorded.

In 1974, Scherk and Schwarz[5] made an important observation that the spin-2 massless particle in the spectrum of vibrational modes of the string should be identified as the graviton. It has been known to the theoretical physicists for a long time that if a good quantum theory of gravity is to be developed then the particle that would carry the gravitational force would have zero mass and two units of spin. Thus the theory of "dual models" was revived under the name "string theory" in the hope of formulating a quantum theory of gravity consistent with the laws of quantum mechanics.

11.5.1 Energy and Length Scales

Hitherto, there is no direct experimental observation that supports the string-like nature of the elementary particles. It is because the present-

[4]Consider a guitar string that has been tuned by stretching the string under tension. Depending on how the string is plucked and how much tension is in the string, different musical notes will be created by the string. These musical notes could be said to be excitation modes of that guitar string under tension.

In a similar manner, in string theory, the elementary particles we observe could be thought of as the "musical notes" or excitation modes of elementary strings.

[5]J. Scherk and J. H. Schwarz, *Nucl. Phys.* **B81**, 118 (1974).

day accelerators have not reached the energy scale at which the string-like nature can be detected. For the applicability of String Theory, we need to discuss the Planckian units to determine the energy scale.

In Relativistic Quantum Mechanics and Quantum Field Theory, we use natural units with $c = 1$ and $\hbar = 1$. In String Theory, one has to use Planckian units with Newton's Gravitational constant $G = 1$ besides $\hbar = 1$ and $c = 1$. This yields Planck length $l_{\rm pl}$, Planck mass $m_{\rm pl}$ and Planck time $t_{\rm pl}$ in terms of G, $\hbar$ and c.

$$\begin{aligned} l_{\rm pl} &= \sqrt{\frac{G\hbar}{c^3}} = 1.61 \times 10^{-33}\,{\rm cm}, \quad m_{\rm pl} = \sqrt{\frac{\hbar c}{G}} = 2.17 \times 10^{-5}{\rm gm}, \\ t_{\rm pl} &= \frac{l_{\rm pl}}{c} = 5.4 \times 10^{-44}\,{\rm sec}. \end{aligned} \tag{11.26}$$

Using Einstein's mass-energy relation $E = mc^2$, the Planck mass $m_{\rm pl} = 10^{19}$ GeV, in energy units. Our present day accelerators have reached the energy only upto 10 TeV $= 10^4$ GeV $= 10^{13}$ eV and the Planck energy is far above. This means that to detect the string-like nature, we need the length scale of 10^{-33} cm or the energy scale of 10^{19} GeV for the accelerators.

It is expected that if the String Theory is a more fundamental unified theory, constructed at 10^{19} GeV, then it must reduce to the Standard Model at 100 GeV.

11.5.2 Basic theory of strings

The basic concepts used in particle physics are extended to the theory of strings[6]. The trajectory of a particle in classical mechanics becomes a world line in relativistic mechanics. Characterizing the world line by a parameter τ, the coordinates of the points on the world line can be represented by $x^\mu(\tau) = \boldsymbol{x}(\tau), t(\tau)$. In Fig. 11.2, we plot a world line on a 2+1 dimensional space-time.

For more than two dimensional space, it is not possible to draw the world line on the plane of a paper but it can be visualized. Given the two space-time points A and B, one can draw very many trajectories connecting the points A and B, but the chosen path is the one which extremises the interval between A and B with the constraint that $\delta(x^\mu) = 0$ at the

[6] A. Ramadevi, *A Pedagogical Introduction to String Theory*, Physics News, **40**, 19 (2010)

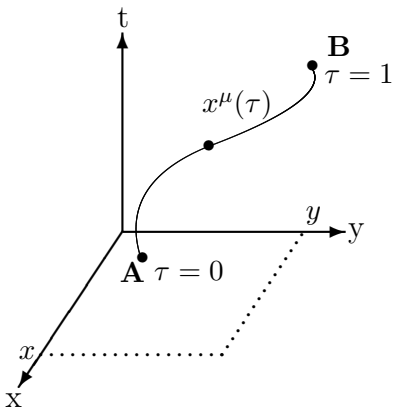


Figure 11.2: World line of a particle parametrized by τ

end points A and B. This will yield a set of second order differential equations for $x^\mu(\tau)$, the solution of which describes the classical path.

11.5.3 Open strings

If we replace the point-like particles by open strings of length l, the trajectories of the string will sweep out a two-dimensional surface called *world sheet* that describes the time evolution of the string. The world sheet can be parametrized by two parameters σ, τ, where $0 \leq \sigma \geq l$ represents the points on the string of length l. The world sheet can be described by the coordinates $x^\mu(\sigma, \tau)$, where the superscript $\mu = 0, 1, 2, \cdots d$ denotes the components of $d + 1$ dimensional space-time coordinate system. In particle physics, we have always assumed $3 + 1$ dimensional space-time but in string theory, the space-time dimensions are determined by the stringent conditions imposed by quantum mechanics and relativity. Such a $d+1$ dimensional space-time is called *the target space* in which the string generates a world sheet.

The open string has two end points $\sigma = 0$ and $\sigma = l$, called the boundary points. To find the classical trajectory of particles, we have extremised the interval between the space-time points A and B. In the case of strings, we need to extremise the area of the world sheet surface to get a set of second order differential equations for $x^\mu(\sigma, \tau)$. In particle theory, we have used the boundary conditions $\delta(x^\mu) = 0$ at A and B. In the case of open strings, we use either the Neumann boundary conditions

at the boundary points

$$\frac{\partial}{\partial\sigma}x^\mu(\sigma,\tau)|_{\sigma=0} = \frac{\partial}{\partial\sigma}x^\mu(\sigma,\tau)|_{\sigma=l} = 0 \tag{11.27}$$

at some directions $\mu = 0, 1, 2, \cdots, p$ where $p \leq d$ and Dirichlet boundary conditions

$$x^m(\sigma = 0, \tau) = b^m, \qquad x^m(\sigma = l, \tau) = c^m, \tag{11.28}$$

(where b_m and c_m are constants) for the remaining directions $m = p + 1, \cdots, d$). There can also be mixed boundary conditions where some directions have Neumann boundary conditions at $\sigma = 0$ and Dirichlet boundary conditions at $\sigma = l$.

Initially, the Dirichlet boundary conditions were considered as unphysical since they were not consistent with the laws of the special theory of relativity in $d + 1$ dimensions and hence strings with only the Newmann boundary conditions were studied. Later, it was pointed out that Nature requires that the laws of special theory of relativity need be satisfied only in 3+1 dimensional space and hence Dirichlet boundary conditions are allowed in other directions. It is found that the Dirichlet boundary conditions lead to non-perturbative spectrum of string theory.

Only the perturbative spectrum involves fundamental strings where the boundary conditions are only Neumann conditions. Extremising the surface area of the world sheet of the open string subjected to the boundary conditions, we obtain the classical solution $x^\mu(\sigma, \tau)$, which can be expanded into Fourier modes a^μ_n that resembles the linear superposition of standing waves. Considering $x^\mu(\sigma, \tau)$ as an operator, quantisation results in commutation relations between the Fourier modes a^μ_n. The modes with positive n resemble harmonic oscillator annihilation operators and those with negative n resemble creation operators $a^\mu_{-n} = (a^\mu_n)^\dagger$.

The non-negative norm imposed by quantum mechanics on the string states and the requirement to satisfy the symmetries of the special theory of relativity results in the 25+1 dimensions for space-time. Thus the space-time dimension emerges from the string theory instead of being assumed as 3+1 in quantum field theory.

The first excited string state in 25+1 dimensional space-time is identified as the massless photon. The vacuum state has negative mass-squared which is usually referred to as *tachyons*.

11.5.4 Closed strings

Unlike open strings, closed strings have no boundary points. The point σ on the closed string coincides with the point $\sigma + l$ since the ends of the string of length l are joined to form a closed loop.

$$x^\mu(\sigma, \tau) = x^\mu(\sigma + l, \tau). \tag{11.29}$$

Again, the classical solution is found by extremising the area of the closed string world sheet subject to the above condition.

It is found that the first excited massless state can be identified as the graviton which is the quantum responsible for gravitational force. Thus the closed string theory incorporates the quantum of gravity. The ground state in closed string theory is again a negative mass-squared state which represents closed string tachyons.

For the study of interactions and scattering, perturbative string theory is developed in a way analogous to perturbation expansion in particle theory using Feynman graphs. In perturbative string theory, Feynman-like diagrams that are used are called *fat graphs*.

11.6 Superstring Theories

The string theory that is developed in 25+1 dimensional space is called *Bosonic String Theory* since all the vibrational modes of the string correspond to integer spins. Generally Bosons represent particles that transmit forces and the Fermions are particles that make up the matter. If you impose *supersymmetry*, then for every Bosonic state, there is a corresponding Fermion and the string theory so obtained is called *Superstring Theory*. There are five kinds of superstring theories, as shown in the table below.

The most important point in formulating a string theory is that it should obey quantum mechanics in a sensible way. For bosonic strings, this is satisfied only if the space-time dimensions number 26. For superstrings we can whittle it down to 10. How we get down to the four space-time dimensions, that we live in, is a different story.

If we ask how to get from ten space-time dimensions to four space-time dimensions, there are many possible ways by which the superfluous six dimensions are made very much smaller than the other four in string theory. This process of compactification of unwanted space-time dimensions yields interesting physics on its own.

Table 11.3: Table of String Theories

Type	Space-time Dimensions	Details
Bosonic	26	Only bosons, no fermions means only forces, no matter, with both open and closed strings. Major flaw: a particle with imaginary mass, called the tachyon
I	10	Supersymmetry between forces and matter, with both open and closed strings, no tachyon, group symmetry is SO(32)
IIA	10	Supersymmetry between forces and matter, with closed strings only, no tachyon, massless fermions spin both ways (nonchiral)
IIB	10	Supersymmetry between forces and matter, with closed strings only, no tachyon, massless fermions only spin one way (chiral)
HO	10	Supersymmetry between forces and matter, with closed strings only, no tachyon, heterotic, meaning right moving and left moving strings differ, group symmetry is SO(32)
HE	10	Supersymmetry between forces and matter, with closed strings only, no tachyon, heterotic, meaning right moving and left moving strings differ, group symmetry is E8 × E8

But the number of string theories has also been shrinking in recent years, because string theorists are discovering that what they thought were completely different theories were in fact different ways of looking at the same theory!

This period in string history has been given the name *the second string revolution*. And now the biggest question is how to collapse the table above into one theory, which some people want to call *the M theory*, for it is *the Mother of all theories*.

At one time, string theorists believed there were five distinct superstring theories: type I, types IIA and IIB, and the two heterotic string theories.

But now it is known that this naive picture was wrong, and that the

five superstring theories are connected to one another as if they are each a special case of a more fundamental M theory. These theories are related by transformations that are called dualities. If two theories are related by a duality transformation, it means that the first theory can be transformed in some way so that it ends up looking just like the second theory. The two theories are then said to be dual to one another under that kind of transformation.

These dualities link quantities that were also thought to be separate. Large and small distance scales, strong and weak coupling strengths – these quantities have always marked very distinct limits of behavior of a physical system, in both classical field theory and quantum particle physics. But strings can obscure the difference between large and small, strong and weak, and this is how these five very different theories end up being related.

Large and small distance

The duality symmetry that obscures our ability to distinguish between large and small distance scales is called T-duality, and comes about from the compactification of extra space dimensions in a ten dimensional superstring theory.

Suppose we are in ten space-time dimensions, which means we have nine space and one time. Take one of those nine space dimensions and make it a circle of radius R, so that traveling in that direction for a distance $L = 2\pi R$ takes you around the circle and brings you back to where you started.

A particle traveling around this circle will have a quantized momentum around the circle, and this will contribute to the total energy of the particle. But a string is very different, because in addition to traveling around the circle, the string can wrap around the circle. The number of times the string winds around the circle is called the winding number, and that is also quantized.

Now the weird thing about string theory is that these momentum modes and the winding modes can be interchanged, as long as we also interchange the radius R of the circle with the quantity $2L_{\text{st}}/R$, where L_{st} is the string length. If R is very much smaller than the string length, then the quantity $2L_{\text{st}}/R$ is going to be very large. So exchanging momentum and winding modes of the string exchanges a large distance scale with a small distance scale.

This type of duality is called T-duality. T-duality relates Type IIA superstring theory to Type IIB superstring theory. That means if we take Type IIA and Type IIB theory and compactify them both on a circle, then switching the momentum and winding modes, and switching the distance scale, changes one theory into the other! The same is also true for the two heterotic theories.

So T-duality obscures the difference between large and small distances. What looks like a very large distance to a momentum mode of a string looks like a very small distance to a winding mode of a string.

Strong and weak coupling

What is a coupling constant? This is some number that tells us how strong an interaction is. A larger coupling constant means a stronger force, and a smaller coupling constant means a weaker force.

For QED, the coupling constant is proportional to the square of the electric charge and the perturbation theory is usually employed to solve any problem. But if the coupling constant is larger, this method of calculation breaks down.

This also can happen in string theory. String theories have a coupling constant. But unlike in particle theories, the string coupling constant is not just a number, but depends on one of the oscillation modes of the string, called the dilaton. Exchanging the dilaton field with minus itself exchanges a very large coupling constant with a very small one. This symmetry is called S-duality. If two string theories are related by S-duality, then one theory with a strong coupling constant is the same as the other theory with weak coupling constant. Since the theory with strong coupling cannot be understood by means of expanding in a series, but the theory with weak coupling can. So if the two theories are related by S-duality, then we just need to understand the weak theory, and that is equivalent to understanding the strong theory.

Superstring theories related by S-duality are: Type I superstring theory with heterotic SO(32) superstring theory, and Type IIB theory with itself.

Starting from any dimension greater than four, it is necessary to consider how these are reduced to four dimensional space-time.

Two different ways have been proposed to resolve this apparent con-

tradiction. The first is to compactify the extra dimensions; i.e., the 6 or 7 extra dimensions are so small as to be undetectable by present day experiments. A standard analogy for this is to consider multidimensional space as a garden hose. If the hose is viewed from a sufficient distance, it appears to have only one dimension, its length. Indeed, think of a ball just small enough to enter the hose. Throwing such a ball inside the hose, the ball would move more or less in one dimension; in any experiment we make by throwing such balls in the hose, the only important movement will be one-dimensional, that is, along the hose. However, as one approaches the hose, one discovers that it contains a second dimension, its circumference. Thus, an ant crawling inside it would move in two dimensions (and a fly flying in it would move in three dimensions). This "extra dimension" is only visible within a relatively close range to the hose, or if one "throws in" small enough objects. Similarly, the extra compact dimensions are only "visible" at extremely small distances, or by experimenting with particles with extremely small wavelengths (of the order of the compact dimension's radius), which in quantum mechanics means very high energies (see wave-particle duality).

T-duality relates the large and small distance scales between string theories, whereas S-duality relates strong and weak coupling strengths between string theories. U-duality links T-duality and S-duality. One such theory is the 11-dimensional M-theory, which requires space-time to have eleven dimensions, as opposed to the usual three spatial dimensions and the fourth dimension of time. The original string theories from the 1980s describe special cases of M-theory where the eleventh dimension is a very small circle or a line, and if these formulations are considered as fundamental, then string theory requires ten dimensions. But the theory also describes universes like ours, with four observable space-time dimensions, as well as universes with up to 10 flat space dimensions, and also cases where the position in some of the dimensions is not described by a real number, but by a completely different type of mathematical quantity. So the notion of space-time dimension is not fixed in string theory: it is best thought of as different in different circumstances.

Appendix A

List of Symbols and Notation

$\sigma_x, \sigma_y, \sigma_z$	Pauli spin matrices $\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},$ $\sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$
$\beta, \alpha_x, \alpha_y, \alpha_z$	Dirac matrices $\beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \quad \boldsymbol{\alpha} \cdot \boldsymbol{p} = \begin{bmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{bmatrix}.$
$\gamma_0, \gamma_x, \gamma_y, \gamma_z$	Gamma matrices $\gamma_0 = \beta, \ \gamma_x = \beta\alpha_x, \ \gamma_y = \beta\alpha_y, \ \gamma_z = \beta\alpha_z.$ $\gamma_0 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}, \quad \boldsymbol{\gamma} = \begin{bmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{bmatrix}.$
γ_5	$\gamma_0\gamma_x\gamma_y\gamma_z$ $\gamma_5 = -i \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$
Scalar product of three-vectors	$\boldsymbol{p} \cdot \boldsymbol{x} = p_x x + p_y y + p_z z;$ $\boldsymbol{A} \cdot \boldsymbol{B} = A_x B_x + A_y B_y + A_z B_z.$
Scalar product of four-vectors	$\mathbf{p} \cdot \mathbf{x} = Et - \boldsymbol{p} \cdot \boldsymbol{x}; \quad \mathbf{A} \cdot \mathbf{B} = A_0 B_0 - \boldsymbol{A} \cdot \boldsymbol{B}.$ The upright bold letters denote four-vectors and the bold italics denote three-vectors.
Contravariant four-vector	$x^\mu, \ \mu = 0, 1, 2, 3.$ $x^0 = ct, \ x^1 = x, \ x^2 = y, \ x^3 = z.$
Covariant four-vector	$x_\mu, \ \mu = 0, 1, 2, 3.$ $x_0 = ct, \ x_1 = -x, \ x_2 = -y, \ x_3 = -z.$

$\not{p}$	$\gamma^\mu p_\mu = \gamma^0 p_0 - \boldsymbol{\gamma}\cdot\boldsymbol{p} = \gamma^0 p_0 - \gamma_x p_x - \gamma_y p_y - \gamma_z p_z$.
$g_{\mu\nu}$	Components of metric tensor $= g^{\mu\nu}$, $\quad \mathrm{diag}(1,-1,-1,-1)$
$\Box$	D'Alembertian operator $\Box = \frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x^\mu} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \boldsymbol{\nabla}^2$.
∂_μ	$\frac{\partial}{\partial x^\mu}$.
∂^μ	$\frac{\partial}{\partial x_\mu}$.
$F^{\mu\nu}$	Electromagnetic field tensor $F^{\mu\nu} = \partial^\nu A^\mu - \partial^\mu A^\nu = \frac{\partial A^\mu}{\partial x_\nu} - \frac{\partial A^\nu}{\partial x_\mu} = -F^{\nu\mu}$.
$[A,B]_-$	$[A,B]_- = AB - BA$. (Commutator bracket)
$\{A,B\}_+$	$\{A,B\}_+ = AB + BA$. (Anticommutator bracket)
$\mathscr{L}$	Lagrangian density.
L	Total Lagrangian $L = \int d^3x \mathscr{L}$.
$\mathscr{A}$	Action $\mathscr{A} = \int dt L$.
$\mathscr{H}$	Hamiltonian density.
H	Total Hamiltonian $H = \int d^3x \mathscr{H}$.

Consider two events at space-time points (x, y, z, t) and $(x + dx, y + dy, z + dz, t + dt)$. Then the interval ds between these two space-time points is defined by

$$ds^2 = c^2 dt^2 - (dx^2 + dy^2 + dz^2). \tag{A.1}$$

The square of the interval ds^2 is the same for all observers in different inertial frames of reference and so is invariant under Lorentz transformation. With this definition, we say that the space-time interval is time-like if

$$ds^2 > 0,$$

and space-like if

$$ds^2 < 0,$$

and light-like if

$$ds^2 = 0.$$

The invariant space-time interval ds is obtained by defining two types of four-vectors, one contravariant four-vector x^μ with an upper index μ and the other called covariant four-vector x_μ with a lower index μ and taking a scalar product (inner product) of them.

Contravariant four-vector

$$x^\mu = x^0, x^1, x^2, x^3 = ct, x, y, z. \tag{A.2}$$

Covariant four-vector

$$x_\mu = x_0, x_1, x_2, x_3 = ct, -x, -y, -z. \tag{A.3}$$

Scalar product

$$ds^2 = \sum_{\mu=0}^{3} dx^\mu dx_\mu = c^2dt^2 - dx^2 - dy^2 - dz^2. \tag{A.4}$$

The relation between the contravariant and covariant vectors can be expressed by defining a metric tensor $g_{\mu\nu}$

$$g_{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}. \tag{A.5}$$

Then

$$\begin{aligned} x_\mu &= \sum_\nu g_{\mu\nu}x^\nu = g_{\mu\nu}x^\nu \\ &= g_{\mu 0}x^0 + g_{\mu 1}x^1 + g_{\mu 2}x^2 + g_{\mu 3}x^3. \end{aligned} \tag{A.6}$$

The repeated index ν implies summation over ν and if this convention is used, then the summation sign is redundant. The relation (A.6) yields

$$x_0 = x^0, \quad x_1 = -x^1, \quad x_2 = -x^2, \quad x_3 = -x^3. \tag{A.7}$$

Since the geometry of the space-time depends on the distribution of matter, leading to the concept of curved space-time, the metric tensor g_{ij} is a function of the space-time coordinates.

The contravariant metric tensor $g^{\mu\nu}$ is given by

$$g^{\mu\nu} = g_{\mu\nu}.$$

A contravariant four-vector A^μ transforms from one coordinate system (unprimed) x^μ to another coordinate system (primed) x'^μ according to the formula

$$A'^\mu = \frac{\partial x'^\mu}{\partial x^\nu} A^\nu; \qquad A^\nu = \frac{\partial x^\nu}{\partial x'^\mu} A'^\mu. \tag{A.8}$$

A covariant four-vector A_μ transforms from the unprimed coordinate system x_μ to the primed coordinate system x'_μ according to the law

$$A'_\mu = \frac{\partial x^\nu}{\partial x'^\mu} A_\nu; \qquad A_\nu = \frac{\partial x'^\mu}{\partial x^\nu} A_\mu. \tag{A.9}$$

If (A^0, A^1, A^2, A^3) are the contravariant components of a four-vector $\mathbf{A}$, then its covariant components A_μ are given by

$$A_\mu = g_{\mu\nu} A^\nu, \qquad (\mu = 0, 1, 2, 3) \tag{A.10}$$

$$A_0 = A^0; \quad A_1 = -A^1; \quad A_2 = -A^2; \quad A_3 = -A^3. \tag{A.11}$$

Thus, we can write the contravariant and covariant components of a four-vector as

$$A^\mu = (A^0, \boldsymbol{A}); \qquad A_\mu = (A^0, -\boldsymbol{A}).$$

The scalar product (inner product) of the two four-vectors $\mathbf{A}$ and $\mathbf{B}$ is

$$\begin{aligned} \mathbf{A} \cdot \mathbf{B} = A^\mu B_\mu &= A^0 B_0 + A^1 B_1 + A^2 B_2 + A^3 B_3 \\ &= A_0 B_0 - A_x B_x - A_y B_y - A_z B_z \\ &= A^0 B_0 - \boldsymbol{A} \cdot \boldsymbol{B}. \end{aligned} \tag{A.12}$$

We adopt the convention of representing four-vectors by upright bold letters and three-component vectors by italic bold letters.

The differential operators in the four-dimensional space-time x^μ transforms as the components of a covariant four-vector.

$$\frac{\partial}{\partial x'^\mu} = \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu}. \tag{A.13}$$

Thus the differential operator is a covariant four-vector.

$$\partial_\nu = \frac{\partial}{\partial x^\nu} = \left(\frac{\partial}{\partial x^0}, \boldsymbol{\nabla} \right). \tag{A.14}$$

The corresponding contravariant four-vector is

$$\partial^\nu = \frac{\partial}{\partial x_\nu} = \left(\frac{\partial}{\partial x_0}, -\boldsymbol{\nabla}\right). \tag{A.15}$$

From the foregoing discussion, it is clear that $g_{\mu\nu}$ contains all the information about the geometry of the space - in this case, Minkowski space-time. If we confine ourselves to the special theory of relativity, the metric tensor $g_{\mu\nu}$ plays a passive role but it plays an active role in general relativity since the space-time geometry is not fixed in advance and can be curved depending on the distribution of matter. So, in special relativity, one can avoid the subtle distinction between the contravariant and covariant vectors and define the scalar product of four-vectors

$$\mathbf{A} \cdot \mathbf{B} = A_0 B_0 - \boldsymbol{A} \cdot \boldsymbol{B} = A_0 B_0 - A_x B_x - A_y B_y - A_z B_z. \tag{A.16}$$

$$\not{p} = \gamma_\mu p_\mu = \gamma_0 p_0 - \boldsymbol{\gamma} \cdot \boldsymbol{p} = \gamma_0 p_0 - \gamma_x p_x - \gamma_y p_y - \gamma_z p_z. \tag{A.17}$$

By the above definition, we can avoid altogether the concept of contravariant and covariant vectors as was done by Feynman in his book on Quantum Electrodynamics (1962). We have followed this convention in chapters 1-5 and only in chapters 6-11, we have introduced the notation of contravariant and covariant vectors.

Appendix B

The Density of Final States

1. One-particle final state

Volume of phase space $= Vd^3p = Vp^2dpd\Omega$.
The number of states dn available for a particle with momentum lying between p and $p + dp$ is obtained by dividing the volume of phase space by the volume of the unit cell h^3.

$$dn = \frac{Vp^2dpd\Omega}{h^3} = \frac{Vp^2dpd\Omega}{(2\pi\hbar)^3}.$$

Let us assume the quantization volume to be unity i.e. $V = 1$ and use natural units i.e. $\hbar = 1,\ c = 1$. The density of states ρ is defined as the number of states available per unit energy interval for a particle scattered within a solid angle $d\Omega$.

$$\rho(E, p) = \frac{dn}{dE\, d\Omega} = \frac{p^2}{(2\pi)^3}\frac{dp}{dE} = \frac{Ep}{(2\pi)^3}. \tag{B.1}$$

The last step is obtained by using the relativistic relation between energy and momentum $E^2 = p^2 + m^2$ such that $dE/dp = p/E$.

2. Two-particle final state

Here we have two particles in the final state with total energy E and total momentum $\boldsymbol{P}$ and the observation is made on the particle 1.

$$E = E_1 + E_2; \qquad \boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2. \tag{B.2}$$

$$\rho(E; p_1, p_2) = \frac{p_1^2}{(2\pi)^3}\frac{dp_1}{dE}. \tag{B.3}$$

We need to evaluate dE/dp_1.

$$\frac{dE}{dp_1} = \frac{dE_1}{dp_1} + \frac{dE_2}{dp_1} = \frac{p_1}{E_1} + \frac{dE_2}{dp_2}\frac{dp_2}{dp_1} = \frac{p_1}{E_1} + \frac{p_2}{E_2}\frac{dp_2}{dp_1}. \tag{B.4}$$

To evaluate dp_2/dp_1, we proceed as follows:

$$\boldsymbol{p}_2 = \boldsymbol{P} - \boldsymbol{p}_1; \qquad p_2^2 = P^2 + p_1^2 - 2\boldsymbol{P}\cdot\boldsymbol{p}_1; \tag{B.5}$$

$$2p_2\frac{dp_2}{dp_1} = 2p_1 - 2\frac{\boldsymbol{P}\cdot\boldsymbol{p}_1}{p_1}$$

$$\frac{dp_2}{dp_1} = \frac{p_1}{p_2} - \frac{\boldsymbol{P}\cdot\boldsymbol{p}_1}{p_1 p_2} = \frac{p_1^2 - \boldsymbol{P}\cdot\boldsymbol{p}_1}{p_1 p_2}. \tag{B.6}$$

Substituting (B.6) into (B.4), we get

$$\begin{aligned}
\frac{dE}{dp_1} &= \frac{p_1}{E_1} + \frac{p_2}{E_2}\frac{p_1^2 - \boldsymbol{P}\cdot\boldsymbol{p}_1}{p_1 p_2} \\
&= \frac{p_1^2 E_2 + (p_1^2 - \boldsymbol{P}\cdot\boldsymbol{p}_1)E_1}{E_1 E_2 p_1} \\
&= \frac{p_1^2(E - E_1) + (p_1^2 - \boldsymbol{P}\cdot\boldsymbol{p}_1)E_1}{E_1 E_2 p_1} \\
&= \frac{p_1^2 E - (\boldsymbol{P}\cdot\boldsymbol{p}_1)E_1}{E_1 E_2 p_1}.
\end{aligned} \tag{B.7}$$

Substituting (B.7) in (B.3), we obtain the two-particle density of states.

$$\rho(E; p_1, p_2) = \frac{E_1 E_2 p_1^2}{(2\pi)^3\{E p_1^2 - E_1(\boldsymbol{P}\cdot\boldsymbol{p}_1)\}}. \tag{B.8}$$

Limiting case
If we allow the mass of the particle 2 to become infinite,

$$m_2 \to \infty, \quad \text{then } E \to \infty \text{ such that } \frac{E}{E_2} \to 1; \quad \frac{E_1}{E_2} \to 0.$$

In this limiting case, the two-particle density of states (B.8)

$$\rho(E; p_1, p_2) \longrightarrow \frac{E_1 p_1}{(2\pi)^3}, \tag{B.9}$$

reduces to the one-particle density of states (B.1).

3. Three-particle final state

Here we have three particles in the final state with total energy E and total momentum $\boldsymbol{P}$ such that

$$E = E_1 + E_2 + E_3; \qquad \boldsymbol{P} = \boldsymbol{p}_1 + \boldsymbol{p}_2 + \boldsymbol{p}_3.$$

The three-particle density of states is a product of one-particle density of states and the other two-particle density of states.

$$\begin{aligned} \rho(E; p_1, p_2, p_3) &= \rho(E_1, p_1)\, \rho(E - E_1; p_2, p_3) \\ &= \frac{E_1 p_1}{(2\pi)^3}\, \frac{E_2 E_3 p_2^3}{(2\pi)^3 \{(E - E_1)p_2^2 - E_2(\boldsymbol{P} - \boldsymbol{p}_1)\cdot \boldsymbol{p}_2\}} \\ &= \frac{1}{(2\pi)^6}\, \frac{E_1 E_2 E_3 p_1 p_2^3}{(E - E_1)p_2^2 - E_2(\boldsymbol{P} - \boldsymbol{p}_1)\cdot \boldsymbol{p}_2}. \end{aligned} \tag{B.10}$$

Limiting case

If $m_3 \to \infty$, then $E_3 \to \infty$. Consequently,

$$\rho(E; p_1, p_2, p_3) = \frac{1}{(2\pi)^6}\, E_1 E_2 p_1 p_2, \tag{B.11}$$

since $\frac{E - E_1}{E_3} \to 1$ and $\frac{E_2}{E_3} \to 0$. The three-particle density of states reduces to a product of two one-particle density of states when one of the particles has an infinite mass.

4. n–particle final state

The construction of three-particle density of states has given us a clue how to construct n–particle density of states. If E is the total energy and $\boldsymbol{P}$, the total momentum of the n–particle state, then the n–particle density of states is the product of one-particle density of states with energy E_1 and momentum p_1 and the (n-1)–particle density of states with energy $E - E_1$ and momentum $\boldsymbol{P} - \boldsymbol{p}_1$.

$$\rho(E; p_1, p_2, \cdots, p_n) = \rho(E_1, p_1)\, \rho(E - E_1; p_2, p_3, \cdots, p_n). \tag{B.12}$$

Thus, the n–particle density of states can be constructed by applying the rule successively.

For additional reading and for a comprehensive account on density of states, the reader is referred to: Alladi Ramakrishnan, *Handbuch der Physik*, Vol. III, p.524, Springer-Verlag (1959); Alladi Ramakrishnan, *Elementary Particles and Cosmic Rays,* p.503, Pergamon Press, Oxford (1962); R.P. Feynman, *Quantum Electrodynamics*, p.81, W.A. Benjamin, New York (1962).

Bibliography

Alladi Ramakrishnan (1962) *Elementary particles and Cosmic Rays*, Pergamon Press, Oxford.

Bjorken, J. D. and Drell, S. D. (1965) *Relativistic Quantum Mechanics*, McGraw Hill, New York.

Bogoliubov, N. N. and Shirkov, D. V. (1959) *Introduction to the Theory of Quantized Fields*, Interscience, New York.

Bromley, D. A. (2000) *Gauge Theory of Weak interactions*, Springer-Verlag, Berlin.

Devanathan, V. (2005) *Quantum Mechanics,* Narosa Publishing House, New Delhi.

Devanathan, V. (1999) *Angular Momentum Techniques in Quantum Mechanics*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Devanathan, V. (2006), *Nuclear Physics, Ch. 14*, Narosa Publishing House, New Delhi.

Feynman, R. P. (1962) *Quantum Electrodynamics*, W. A. Benjamin, Inc., New York.

Feynman, R. P. (1990), *QED, The Strange Theory of Light and Matter*, Pemguin.

Greiner, Walter (1990) *Relativistic Quantum Mechanics*, Springer-Verlag, Berlin.

Greiner, W and Schäfer, A (1994), *Quantum Chromodynamics*, Springer-Verlag, Berlin.

Griffiths, D. (1987) *Introduction to Elementary Particle Physics*, Wiley, New York.

Jauch, J. M. and Rohrlich, F. (1955), *The Theory of Photons and Electrons*, Addison Wesley Publishing Company, Massachusetts, U.S.A.

Lahiri, A and Pal, P. B. (2005) *A First Book of Quantum Field Theory, Second Edition*, Narosa Publishing House, New Delhi.

Lee, T. D. (1981) *Particle Physics and Introduction to Field Theory*, Harwood Academic, New York.

Mandl, F and Shaw, G (1984) *Quantum Field Theory*, John Wiley and Sons Ltd.

Parthasarathy, R (2010) Relativistic Quantum Mechanics, Narosa Publishing House, New Delhi.

Peshkin, M. E. and Schroeder, D. V. (1995) *An Introduction to Quantum Field Theory*, Addison-Wesley, Massachusetts.

Rajesh Gopakumar (2010) "What is the AdS/CFT Correspondence?", *Physics News.* **40**, 34-44.

Ramadevi, P. (2010) "String Theory - A Pedagogical Review', *Physics News.* **40**, 19-33.

Roy, D. P. (1999), "Basic Constituents of Matter and their Interactions - A Progress Report", *TIFR/TH/99-59; arXiv:hep-ph/9912523.*

Ryder, Lewis H. (1985) *Quantum Field Theory*, Cambridge University Press.

Schweber, S. S. (1994) *QED and the Men who made it: Dyson, Feynman, Schwinger and Tomanaga*, Princeton University Press, Princeton.

Schweber, S. S., Bethe, H. A. and Hoffmann, F. D. (1956), *Mesons and Fields, Vol. 1*, Row, Peterson and Company, New York.

Schwinger, J (Ed.) (1958), *Selected Papers on Quantum Electrodynamics*, Dover Publications, New York.

Thangappan, V. K. (1985), *Quantum Mechanics*, Wiley Eastern Limited, India.

Venkataraman, G. (1994), *Quantum Revolution II; QED: The Jewel of Physics*, University Press India) Ltd. Hyderabad, India.

Weinberg, Steven, (1995) *The Quantum Theory of Fields, 3 Vols.*, Cambridge University Press.

Zee, A. (2003) *Quantum Field Theory in a Nutshell*, Princeton University Press, Princeton.

Subject Index

Action integral, 127
Anomalous magnetic moment of electron, 13
Anti-commutation relations, 14, 129
Anti-particle, 10

Basic vertices in QED, 56–58
Beta decay interaction, 218
Bhabha scattering, 80
Bremsstrahlung, 87–91
  Feynman diagrams, 87
  soft photon limit, 90

Canonical commutation relations, 1
Canonical conjugate variables, 1
Classical fields, 126
Classical mechanics, 1, 124–126
  A brief review, 124–126
  Euler-lagrange equations of motion, 125
  Hamilton action principle, 124
  Hamiltonian formalism, 124
  Hamilton's canonical equations of motion, 126
  Lagrangian formalism, 124
  the variational principle, 124
Commutation relations, 14, 129
Complex scalar fields, 148
  charge and current density, 150
  Lagrangian density, 148
  particles and antiparticles, 152
Compton scattering, 65–74
  Compton shift in frequency, 67
  differential cross section, 73
  evaluation of traces, 70
  extreme relativistic limit, 74
  Feynman diagrams, 68
  Klein-Nishina formula, 72
  non-relativistic limit, 73
  Thomson scattering formula, 74
  total cross section, 73
  transition probability, 70
Conjugate momentum field, 14
Conservation of baryon numbers, 233
Conservation of lepton numbers, 233
Continuity equation, 2, 4, 8–9
Contravariant and covariant vectors, 243, 245
Covariant commutation relations, 153

D'Alembertian operator, 183, 244
Density of final states, 72, 80, 86, 248–250
Dimensional regularization, 221
Dirac equation, 5, 11, 21–37
  free particle solutions, 9–10, 21–37
  orthogonal and closure properties, 24
  projection operators, 26–27, 30
  in Feynman's notation, 27
  Lorentz boost operator, 25, 31
  negative eigenvalue states, 29–30
  negative energy solutions, 23–24
  plane wave solutions, 21–37, 164
  positive eigenvalue states, 28–29
  positive energy solutions, 23–24

sum over spin states, 26, 33
Dirac field, 5, 163–178
annihilation and creation operators, 168
anticommutation relations, 163, 171
conjugate momentum density, 167
covariant anticommutation relations, 172
Fourier decomposition, 167
Hamiltonian density, 167
Lagrangian density, 166
quantization, 169
total charge, 168
total energy, 168
Dirac Hamiltonian, 6
Dirac matrices, 5–8
Dirac spinors, 9
Dirac vacuum, 10
Dirac's hole theory, 10
Dyson's chronological operator, 203

Electro-weak theory, 229
Electromagnetic field, 179–196
Gupta-Bleuler formalism, 188, 189
Indefinite metric operator η, 190
Lagrangian density, 185
Lorentz condition, 181, 191
Lorentz gauge, 193
Maxwell's equations, 180
quantization, 186
scalar and vector potentials, 180
Electromagnetic field tensors, 183
Electron self energy, 109–114, 116, 215
bare and physical mass, 114
mass renormalization, 114
shift in mass, 114
Electron spin, 19
gyromagnetic ratio, 19
Electron-electron scattering, 74–80
evaluation of matrix elements, 76
Feynman diagrams, 74
in c.m. frame, 79
matrix elements, 74
Electron-positron scattering, 80–83
Feynman diagrams, 81
in c.m. frame, 82
Euler-Lagrange equation of motion, 125

Fermi's golden rule, 34, 64, 96
Feynman diagrams, 11, 53–58, 211
electron self energy, 110, 116, 216
in momentum representation, 53
in space-time representation, 53
photon self energy, 216
vacuum diagram, 216
vacuum polarization, 216
Feynman's positron theory, 11–12
Feynman's propagation kernel, 48
in momentum representation, 49
First quantization, 1, 123, 130
Four vectors, 32, 243–245
scalar product, 243, 245, 247

Gamma matrices, 28
algebra, 35–37
Evaluation of traces, 37
traces, 36
Gauge Bosons, 223
Gauge covariant derivative, 223
Gauge fields, 223
Gauge group SU(3) $\times$ SU(2) $\times$ U(1), 228, 231
Gauge theories, 221–227
Gauge theory of electromagnetic interaction, 221–223
Gravitational interaction, 233
Gravitons, 233
Gupta-Bleuler formalism, 187, 189

Heisenberg's uncertainty relation, 1
Higgs Boson, 226, 230
Higher order corrections, 114–117

Infra-red slavery, 229
Interaction between fields, 197–218
Intermediate vector Bosons $W^{\pm}, Z^0$, 228, 232

Invariant perturbation theory, 199–202

Klein-Gordon Equation, 3
Klein-Gordon field, 142
Klein-Nishina formula, 72

Lamb shift, 13, 114

Metric tensor $g^{\mu\nu}$, 245
Minkowski space, 127, 247
Mott scattering cross section, 64
Møller scattering, 80
Muon pair production, 92–96
  application to hadron production, 96-97
  Feynman diagrams, 93
  total cross section, 96

Neutrino oscillations, 233
Noether's theorem, 232
Non-conservative forces, 128
Normal ordering, 147–148, 170
Normal product, 204–213

Pair annihilation, 83–87
  density of final states, 86
  Feynman diagrams, 83
Pair production, 91–92
  Feynman diagrams, 92
Pauli-Villars regularization, 221
Photon self energy, 117, 215
Poisson bracket, 126
Positron scattering by external field, 213
Principle of least action, 127
Probability current density, 2, 4, 8
Probability density, 2, 4, 9
Propagation kernel, 44–50
  differential equation, 46
  Feynman's, 48
  in Dirac theory, 47
  in momentum representation, 49
  in Schrödinger theory, 44
  integral equation, 47

Quantization of the field, 129–135
Quantum chromodynamics, 230
Quantum Electrodynamics, 12–13, 60–109
  Bremsstrahlung, 87–91
  Compton scattering, 65–74
  Electron-electron scattering, 74–80
  Electron-positron scattering, 80–83
  Muon pair production, 92–96
  Pair annihilation, 83–87
  Pair production, 91–92
  Rutherford scattering, 60–65
Quantum Field Theory, 13–15, 123–242
  Dirac field, 163–178
  Electromagnetic field, 179–196
  Interacting fields, 197–218
  Scalar fields, 142–155
  Schrödinger field, 130–135
Quarks and leptons, 229

Radiation corrections, 109–117
Regularization and Renormalization, 15, 110, 220
Relativistic energy-momentum relation, 3
Relativistic fields, 135–137
Renormalization programs, 219, 220
Rutherford scattering, 60–65
  Mott scattering cross section, 64
Rutherford scattering formula, 64

S-matrix, 15, 202
  First order terms, 209
  Second order terms, 213
S-matrix expansion, 209
  reduction into normal products, 203
Scalar fields, 142–155
  annihilation and creation operators, 146

charge-current density, 150
complex scalar fields, 148
covariant commutation relations, 153
Fourier decomposition of the field, 143
ground state and normal ordering, 147
number operator, 146
one-component real field, 142
particles and anti-particles, 152
propagator, 154–155
quantization, 146
Schrödinger Equation, 2
Schrödinger field, 130–135
annihilation and creation operators, 131
Lagrangian density, 130
quantization, 131
quantization into Bosons, 132
quantization into Fermions, 133
Second quantization, 14, 129
Space-time interval, 244
space-like, 244
time-like, 244
Spontaneous symmetry breaking,
Standard model, 228–233
Theoretical framework, 231
String theory, 234–242
closed strings, 238
open strings, 236
Superstring theories, 238–242
large and small distances, 240
S-duality, 241
strong and weak coupling, 241
T-duality, 241
Symmetries
global symmetry, 221
local symmetry, 222

Transition amplitude, 51
First order matrix elements, 51
Second order matrix elements, 52
Tomanaga-Schwinger equation, 198
Transition Probability, 34
Fermi's golden rule, 34

Vacuum diagram, 215
Vacuum polarization, 117, 215
Vertex correction, 117

Wick's chronological product, 205
Wick's contractions, 206
Wick's theorem, 208
World line, 12, 208